

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
25. Juli 2002 (25.07.2002)

PCT

(10) Internationale Veröffentlichungsnummer  
**WO 02/057464 A2**

- (51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/82 (74) Anwalt: GOLDSCHIED, Bettina; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).
- (21) Internationales Aktenzeichen: PCT/EP02/00461
- (22) Internationales Anmeldedatum:  
18. Januar 2002 (18.01.2002)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:  
101 02 338.3 19. Januar 2001 (19.01.2001) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF PLANT SCIENCE GMBH [DE/DE]; 67056 Ludwigshafen (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): LERCHL, Jens [DE/SE]; Onsjövägen 17, S-26831 Svalöv (SE). DUWENIG, Elke [DE/DE]; Schanzstraße 46, 67063 Ludwigshafen (DE). BISCHOFF, Friedrich [DE/DE]; Albinstrasse 11, 55116 Mainz (DE). HEINZ, Ernst [DE/DE]; Püttkampsweg 13, 22609 Hamburg (DE). DREXLER, Hjärdis [DE/DE]; Mendelssohnstr. 33, 22761 Hamburg (DE). SCHEFFLER, Jodi [US/US]; 51 County Road 228, Oxford, MS 38655 (US).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI, GM, GR, GU, HK, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Veröffentlicht:  
— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts
- Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHOD FOR THE EXPRESSION OF BIOSYNTHETIC GENES IN PLANT SEEDS USING NOVEL MULTIPLE EXPRESSION CONSTRUCTS

(54) Bezeichnung: VERFAHREN ZUR EXPRESSION VON BIOSYNTHESEGENEN IN PFLANZLICHEN SAMEN UNTER VERWENDUNG VON NEUEN MULTIPLEN EXPRESSIONSKONSTRUKTEN

(57) Abstract: The invention relates to expression cassettes, combinations thereof and vectors containing said expression cassettes, containing plant promoters with an expression specificity for plant seeds, in particular linseed and the use of said expression cassettes or vectors for the recombinant expression of heterologous genes in plants. The invention further relates to transgenic plants, transformed by means of said expression cassettes, or vectors, cultures, parts or transgenic propagations derived therefrom and the use of the above as foodstuff, animal feedstuff, seedstuff, pharmaceuticals, fine chemicals or industrial raw material.

(57) Zusammenfassung: Die Erfindung betrifft Expressionskassetten, deren Kombination und Vektoren enthaltend die Expressionskassetten, die pflanzliche Promotoren mit einer Expressionsspezifität für pflanzliche Samen insbesondere Leinsamen enthalten, sowie die Verwendung dieser Expressionskassetten oder Vektoren zur rekombinanten Expression heterologer Gene in Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder transgenes Vermehrungsgut, sowie die Verwendung derselben als Nahrungs-, Futtermittel oder Saatgut, Pharmazeutika, Feinchemikalien oder als industrieller Grundstoff.

WO 02/057464 A2

Verfahren zur Expression von Biosynthesegenen in pflanzlichen Samen unter Verwendung von neuen multiplen Expressionskonstrukten

## 5 Beschreibung

Die Erfindung betrifft Expressionskassetten, deren Kombination und Vektoren enthaltend die Expressionskassetten, die pflanzliche Promotoren mit einer Expressionsspezifität für pflanzliche Samen  
10 insbesondere Leinsamen enthalten, sowie die Verwendung dieser Expressionskassetten oder Vektoren zur rekombinanten Expression heterologer Gene in Pflanzen. Die Erfindung betrifft ferner mit diesen Expressionskassetten oder Vektoren transformierte transgene Pflanzen, davon abgeleitete Kulturen, Teile oder transgenes  
15 Vermehrungsgut, sowie die Verwendung derselben als Nahrungs-, Futtermittel oder Saatgut, Pharmazeutika, Feinchemikalien oder als industrieller Grundstoff.

Die erfindungsgemäßen Expressionskassetten werden vorteilhaft in  
20 einem Verfahren zur Herstellung von ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen und/oder ein Verfahren zur Herstellung von Triglyceriden mit erhöhtem Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen verwendet. Im Verfahren finden vorteilhaft die Nukleinsäuresequenzen  
25 SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder 11 Anwendung, die unter Verwendung der Expressionskassetten exprimiert werden. Diese vorgenannten Nukleinsäuren sind im Verfahren sowie zur Herstellung eines transgenen Organismus bevorzugt einer transgenen Pflanze oder eines transgenen Mikroorganismus mit erhöhtem Gehalt  
30 an Fettsäuren, Ölen oder Lipiden mit ungesättigten C<sub>18</sub>-, C<sub>20</sub>-, oder C<sub>22</sub>-Fettsäuren geeignet. Außerdem können mit Hilfe der erfindungsgemäßen Expressionskassetten weitere Gene neben den Nukleinsäuresequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 und SEQ ID NO: 11 oder seine Homologen, Derivate oder Analoga sowie  
35 Genkonstrukte, die diese Gene oder ihre Homologen, Derivate oder Analoga umfasst, sowie ihre Verwendung allein oder in Kombination mit weiteren Biosynthesegenen bevorzugt Biosynthesegene für polyungesättigter Fettsäuren, wie vorteilhaft in SEQ ID NO: 7 und SEQ ID NO: 9 dargestellt, in Organismen bevorzugt Pflanzen exprimiert  
40 werden.

Eine Reihe von Produkten und Nebenprodukten natürlich vorkommender Stoffwechselprozesse in Mikroorganismen, tierischen und pflanzlichen Zellen sind für viele Industriezweige, ein-  
45 schließlich der Futtermittel-, Nahrungsmittel-, Kosmetik- und pharmazeutischen Industrie, nützlich. Zu diesen gemeinsam als "Feinchemikalien" bezeichneten Molekülen gehören beispielsweise

## 2

Lipide und Fettsäuren, unter denen eine beispielhafte Klasse die mehrfach ungesättigten Fettsäuren sind. Fettsäuren und Triglyceride haben eine Vielzahl von Anwendungen in der Lebensmittelindustrie, der Tierernährung, der Kosmetik und im Pharmabereich. Je nachdem ob es sich um freie gesättigte oder ungesättigte Fettsäuren oder um Triglyceride mit einem erhöhten Gehalt an gesättigten oder ungesättigten Fettsäuren handelt, sind sie für die unterschiedlichsten Anwendungen geeignet, so werden beispielsweise mehrfach ungesättigte Fettsäuren (polyunsaturated fatty acids, PUFAs) Babynahrung zur Erhöhung des Nährwertes zugesetzt. PUFAs haben weiterhin einen positiven Einfluss auf den Cholesterinspiegel im Blut von Menschen und eignen sich daher zum Schutz gegen Herzkrankheiten. So finden sie in verschiedenen diätischen Lebensmitteln oder Medikamenten Anwendung.

Besonders geeignete Mikroorganismen zur Herstellung von PUFAs sind Mikroorganismen wie Thraustochytrien oder Schizochytrien-Stämme, Algen wie Phaeodactylum tricornutum oder Cryptocodinium-Arten, Ciliaten, wie Stylonychia oder Colpidium, Pilze, wie Mortierella, Entomophthora oder Mucor. Durch Stammselektion ist eine Anzahl von Mutantenstämmen der entsprechenden Mikroorganismen entwickelt worden, die eine Reihe wünschenswerter Verbindungen, einschließlich PUFAs, produzieren. Die Selektion von Stämmen mit verbesserter Produktion eines bestimmten Moleküls ist jedoch ein zeitraubendes und schwieriges Verfahren.

Alternativ kann die Produktion von Feinchemikalien geeigneterweise über die Produktion von Pflanzen, die so entwickelt sind, dass sie die vorstehend genannten PUFAs herstellen, im großen Maßstab durchgeführt werden. Besonders gut für diesen Zweck geeignete Pflanzen sind Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten wie Raps, Canola, Lein, Soja, Sonnenblumen, Borretsch und Nachtkerze. Aber auch andere Nutzpflanzen, die Öle oder Lipide und Fettsäuren enthalten, sind gut geeignet, wie in der eingehenden Beschreibung dieser Erfindung erwähnt. Mittels herkömmlicher Züchtung ist eine Reihe von Mutantpflanzen entwickelt worden, die ein Spektrum an wünschenswerten Lipiden und Fettsäuren, Cofaktoren und Enzymen produzieren. Die Selektion neuer Pflanzensorten mit verbesserter Produktion eines bestimmten Moleküls ist jedoch ein zeitaufwändiges und schwieriges Verfahren oder sogar unmöglich, wenn die Verbindung in der entsprechenden Pflanze nicht natürlich vorkommt, wie im Fall von mehrfach ungesättigten C<sub>18</sub>-, C<sub>20</sub>-Fettsäuren und C<sub>22</sub>-Fettsäuren und solchen mit längeren Kohlenstoffketten.

## 3

Aufgrund der positiven Eigenschaften ungesättigter Fettsäuren hat es in der Vergangenheit nicht an Ansätzen gefehlt, Gene, die an der Synthese von Fettsäuren bzw. Triglyceriden beteiligt sind, für die Herstellung von Ölen in verschiedenen Organismen mit geändertem Gehalt an ungesättigten Fettsäuren verfügbar zu machen. So wird in WO 91/13972 und seinem US-Äquivalent eine  $\Delta$ -9-Desaturase beschrieben. In WO 93/11245 wird eine  $\Delta$ -15-Desaturase in WO 94/11516 wird eine  $\Delta$ -12-Desaturase beansprucht.  $\Delta$ -6-Desaturasen werden in WO 93/06712, US 5,614,393, WO 96/21022 und WO 99/27111 beschrieben. Weitere Desaturasen werden beispielsweise in EP-A-0 550 162, WO 94/18337, WO 97/30582, WO 97/21340, WO 95/18222, EP-A-0 794 250, Stukey et al., J. Biol. Chem., 265, 1990: 20144-20149, Wada et al., Nature 347, 1990: 200-203 oder Huang et al., Lipids 34, 1999: 649-659 beschrieben. In WO 96/13591 wird eine  $\Delta$ -6-Palmitoyl-ACP-Desaturase beschrieben und beansprucht. Die biochemische Charakterisierung der verschiedenen Desaturasen ist jedoch bisher nur unzureichend erfolgt, da die Enzyme als membrangebundene Proteine nur sehr schwer zu isolieren und charakterisieren sind (McKeon et al., Methods in Enzymol. 71, 1981: 12141-12147, Wang et al., Plant Physiol. Biochem., 26, 1988: 777-792).

In Hefen konnte sowohl eine Verschiebung des Fettsäurespektrums zu ungesättigten Fettsäuren hin als auch eine Steigerung der Produktivität nachgewiesen werden (siehe Huang et al., Lipids 34, 1999: 649-659, Napier et al., Biochem. J., Vol. 330, 1998: 611-614). Die Expression der verschiedenen Desaturasen in transgenen Pflanzen zeigte allerdings nicht den gewünschten Erfolg. Eine Verschiebung des Fettsäurespektrums zu ungesättigten Fettsäuren hin konnte gezeigt werden, gleichzeitig zeigte sich aber, dass die Syntheseleistung der transgenen Pflanzen stark nachließ, das heißt gegenüber den Ausgangspflanzen konnten nur geringere Mengen an Ölen isoliert werden.

Weder in Hefen noch in Pflanzen werden natürlicherweise mehrfach ungesättigte C<sub>20</sub>- und/oder C<sub>22</sub>-Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül wie Arachidonsäure (ARA) und/oder Eicosapentaensäure (EPA) und/oder Docosahexaensäure (DHA) hergestellt.

40

Nach wie vor besteht daher ein großer Bedarf an neuen Genen, die für Enzyme kodieren, die an der Biosynthese ungesättigter Fettsäuren beteiligt sind und es ermöglichen, diese in einem technischen Maßstab herzustellen. Keines der bisher bekannten biotechnologischen Verfahren zur Herstellung von mehrfach

ungesättigten Fettsäuren liefert die vorgenannten Fettsäuren in wirtschaftlich nutzbaren Mengen.

Bei der Expression von Genen in Pflanzen gibt es immer wieder  
5 Probleme, dass heißt es kommt durch die Expression nicht zur erwarteten Steigerung bei der Herstellung des gewünschten Wertprodukts.

Verschiedene Methoden zum Einschleusen von Genen in das Genom von  
10 Pflanzen sind bekannt (Halford NG, Shewry PR, Br. Med. Bull., 2000; 56: 62-73). Ziel ist die Herstellung von Pflanzen mit vorteilhaften, neuen Eigenschaften, zum Beispiel zur Steigerung der landwirtschaftlichen Produktivität, zur Qualitätssteigerung bei Nahrungsmitteln oder zur Produktion bestimmter Chemikalien oder  
15 Pharmazeutika (Dunwell JM, J. Exp. Bot., 2000: 51 Spec No: 487-96).

Eine Grundvoraussetzung für die transgene Expression bestimmter Gene ist die Bereitstellung pflanzenspezifischer Promotoren. Pro-  
20 motoren sind wichtige Werkzeuge in der Pflanzenbiotechnologie, um die Expression eines bestimmten Gens in einer transgenen Pflanze lokal und zeitlich zu steuern und so bestimmte Wesensmerkmale der Pflanze auszunutzen bzw. erst zu erzielen. Verschiedene Promotoren für diverse Pflanzenarten, bestimmte Pflanzengewebe und Ent-  
25 wicklungsstadien sind bekannt.

Verwendet werden zum Beispiel konstitutive Promotoren wie der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpro-  
motor oder der Promotor des 35S-Transkriptes des Blumenkohlmo-  
30 saikvirus (CaMV) (Odell et al., Nature 1985: 313,810-812), der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al., Plant Mol. Biol. 1995, 29:637-646), die Promotoren der vakuolärer ATPase Untereinheiten oder der Pro-  
motor eines prolinreichen Proteins aus Weizen (WO 91/13991).  
35 Nachteilig bei diesen Promotoren ist, dass sie in fast allen Geweben der Pflanze konstitutiv aktiv sind. Eine gezielte Expression von Genen in bestimmten Pflanzenteilen oder zu bestimmten Entwicklungszeitpunkten ist mit diesen Promotoren nicht möglich.

Promotoren, deren Aktivität gewebsspezifisch oder entwicklungsab-  
40 hängig reguliert ist, wurden isoliert. Spezifitäten sind für die Antheren, Ovarien, Blüten, Blätter, Stengel, Wurzeln und Samen beschrieben. Die Stringenz der Spezifität, als auch die Expressionsaktivität dieser Promotoren ist sehr unterschiedlich. Zu  
45 nennen sind Promotoren, die eine blattspezifische Expression gewährleisten, wie der Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubi-

## 5

sco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al., EMBO J. 8 (1989), 2445 - 245).

5 Weitere Promotoren sind beispielsweise spezifische Promotoren für Knollen, Speicherwurzeln oder Wurzeln, wie beispielsweise der Patatin Promotor Klasse I (B33), der Promotor des Cathepsin D Inhibitors aus Kartoffel, der Promotor der Stärke Synthese (GBSS1) oder der Sporamin Promotor, fruchtspezifische Promotoren, wie  
10 beispielsweise der fruchtspezifische Promotor aus Tomate (EP-A 409625), fruchtreifungsspezifische Promotoren, wie beispielsweise der fruchtreifungsspezifische Promotor aus Tomate (WO 94/21794), blütenspezifische Promotoren, wie beispielsweise der Phytoen-Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO  
15 98/22593).

Eine Variation der Aktivität eines Promotoren anhängig vom Entwicklungsstadium der Pflanze wird unter anderem von Baerson et al. beschrieben (Baerson SR, Lamppa GK. Plant Mol Biol.  
20 1993;22(2):255-67).

Samenspezifische Promotoren sind aufgrund der Bedeutung des Samens als eine der Hauptnahrungs- bzw. Futterquellen von Mensch und Tier und als Produktionsort für Wertstoffe von besonderen Interesse. Bekannt sind Promotoren, die eine Expression in Samen und pflanzlichen Embryonen steuern. So wurden beispielsweise die Promotoren von Genen identifiziert, die für Speicherproteine verschiedener Dicotyledonen kodieren. Samenspezifische Promotoren sind zum Beispiel der Promotor des Phaseolins (US 5504200, Bustos  
25 MM et al., Plant Cell. 1989;1(9):839-53), des 2S Albumingens (Joseffson LG et al., J Biol Chem 1987, 262:12196-12201), des Legumins (Shirsat A et al., Mol Gen Genet. 1989;215(2):326-331), des USP (unknown seed protein; Bäumlein H et al., Molecular & General Genetics 1991, 225(3):459-67) des Napin Gens (Stalberg K, et  
30 al., L. Planta 1996, 199:515-519), des Saccharosebindepoteins (WO 00/26388) oder der LeB4-Promotor (Bäumlein H et al., Mol Gen genet 1991, 225: 121-128). Diese steuern eine hohe samenspezifische Expression von Speicherproteinen.

40 Trotz der allgemeinen Annahme, daß pflanzliche Promotoren von einer Spezies auf die andere Übertragbar sind und auch in artfremden Pflanzenspezies ähnliche Aktivitäten und Spezifitäten aufweisen, mehrten sich Hinweise auf Einschränkungen von dieser Annahme. So zeigte es sich, daß die Höhe der transgenen Expression von heterologen Genen unter Kontrolle dieser Promotoren, oft stark von  
45 der Art der Wirtspflanze abhängig ist. Es wurde festgestellt, dass die Expression nicht immer absolut zelltypspezifisch ist.

## 6

Unterschiede im Expressionsmuster und der Expressionsstärke eines bestimmten Promotors können durch unterschiedliche Wirtspflanzen oder durch unterschiedliche Insertionsorte in das Genom der Wirtspflanze bedingt sein (Goossens A et al., Plant Phys 1999, 5 120:1095-1104).

Durch eine Genexpression in anderen Pflanzenteilen kann die Verwendung eines Promoters in einer anderen Pflanzenart sehr eingeschränkt sein. Etwa wenn die Expression des Genes in den Metabolismus der Zelle, die Zusammensetzung der Membranlipide oder die Biosynthese eingreift. 10

Es bestand daher die Aufgabe weitere Expressionskassetten für die Expression in Pflanzen zur Verfügung zu stellen. Und diese für 15 die Expression von Genen vorteilhaft Biosynthesegenen allein oder gegebenenfalls in Kombination mit anderen Enzymen in einem Verfahren beispielsweise zur Herstellung mehrfach ungesättigter Fettsäuren zu verwenden. Diese Aufgabe wurde durch die erfindungsgemäßen Expressionskassette mit einer Struktur ausgewählt 20 aus der Gruppe gelöst:

- a) L1 - Promotor - Strukturgen - L2,
- b) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Struktur- 25 gen - L2,
- c) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2,

30 wobei L1, L2, Promotor und Strukturgen die folgende Bedeutung hat:

L1 = SEQ ID NO: 32 oder eine äquivalente Restriktionsschnittstellen enthaltende Sequenz, 35

L2 = unabhängig voneinander SEQ ID NO: 33, 34 oder 35 oder äquivalente Restriktionsschnittstellen enthaltende Sequenzen,

Promotor = pflanzlicher Promotor 40

Strukturgen = eine in Pflanzen exprimierbare Nukleinsäuresequenz.

Vorteilhaft ist das Strukturgen ein Biosynthesegen, bevorzugt 45 ist es ein Biosynthesegen des Lipid- oder Fettsäurestoffwechsels, vorteilhaft ein pflanzliches Gen. In einer besonders vorteilhaften Ausführungsform ist das Strukturgen eine Nukleinsäuresequenz,

## 7

die für Proteine ausgewählt aus der Gruppe :

Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-  
 5 Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxxygenasen, Triacylglycerol-Lipaser, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n), kodiert.

10 Ganz besonders bevorzugt ist das Strukturgen eine Nukleinsäuresequenz ausgewählt aus der Gruppe:

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,  
 15
- b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,  
 20
- c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen,  
 25 ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

Unter einer äquivalente Restriktionsschnittstellen enthaltende  
 30 Sequenz im Sinne der Erfindung sind Sequenzen zu verstehen, die Restriktionsschnittstellen enthalten, die für den Aufbau multiplexer Expressionskassetten geeignet sind, das heißt geeigneter Weise nicht im Strukturgen vorhanden sind oder im binären Vektor. Solche Restriktionsschnittstellen wie beispielhaft EcoRI, BamHI,  
 35 SacI, PstI, NcoI, NdeI, BglI, BglII, XhoI, Xba und weitere sind dem Fachmann bekannt und können einschlägigen Fachbüchern entnommen werden.

Die erfindungsgemäßen Expressionskassetten können für die Expression von Genen in wirtschaftlich wichtigen Kulturpflanzen wie  
 40 beispielsweise Lein, für die keine endogenen samenspezifischen Promotoren bekannt waren, verwendet werden. Lein ist, wie die vorliegenden Arbeiten zeigten, für eine samenspezifische Expression von Genen besonders problematisch, da offensichtlich mehrere  
 45 Promotoren, die vom Fachmann für samenspezifische Expression in anderen Pflanzen routinemäßig benutzt werden, in z.B. in anderen Pflanzen wie Lein nicht funktionieren, das heißt nicht zu einer

## 8

Transkription bzw. letztlich zu einer Expression der mRNA des Strukturgens führt.

Verwendung der oben genannten erfindungsgemäßen Expressionskas-  
5 setten in einem Verfahren zur Herstellung von Fettsäureestern mit einem erhöhtem Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen dadurch gekennzeichnet, daß man mindestens eine Nukleinsäuresequenz ausgewählt aus der Gruppe

- 10 a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,
- b) Nukleinsäuresequenzen, die aufgrund des degenerierten  
15 genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,
- c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5  
20 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne dass die enzymatische Wirkung  
25 der Polypeptide wesentlich reduziert ist,

in einen Fettsäureester produzierenden Organismus einbringt, anzieht und die dem Organismus enthaltenden Fettsäureester isoliert.

- 30 Bei den im Verfahren verwendeten Nukleinsäuresequenzen handelt es sich um isolierte Nukleinsäuresequenzen, die für Polypeptide mit  $\Delta$ -5-,  $\Delta$ -6- oder  $\Delta$ -12-Desaturaseaktivität codieren.

- 35 Im Verfahren werden vorteilhaft Fettsäureester mit mehrfach ungesättigten  $C_{18}$ -,  $C_{20}$ - und/oder  $C_{22}$ -Fettsäuremolekülen mit mindestens zwei Doppelbindungen im Fettsäureester hergestellt. Bevorzugt enthalten diese Fettsäuremoleküle drei, vier oder fünf Doppelbindungen und führen vorteilhaft zur Synthese von Arachidonsäure (ARA), Eicosapentaensäure (EPA) oder Docosahexaensäure  
40 (DHA).

- Die Fettsäureester mit mehrfach ungesättigten  $C_{18}$ -,  $C_{20}$ - und/oder  $C_{22}$ -Fettsäuremolekülen können aus den Organismen, die für die Her-  
45 stellung der Fettsäureester verwendet wurden, in Form eines Öls oder Lipids beispielsweise in Form von Verbindungen wie Sphingolipide, Phosphoglyceride, Lipide, Glycolipide, Phospholipide,

Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Fettsäureester, die die mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen enthalten, isoliert werden.

5

Als Organismus für die Herstellung im Verfahren kommen prinzipiell alle prokaryontischen oder eukaryontischen Organismen wie prokaryontische oder eukaryontische Mikroorganismen wie gram-positive oder gram-negative Bakterien, Pilze, Hefen, Algen, Ciliaten, tierische oder pflanzliche Zellen, Tiere oder Pflanzen wie Moose, zweikeimblättrige oder einkeimblättrige Pflanzen in Frage. Vorteilhaft werden Organismen im erfindungsgemäßen Verfahren verwendet, die zu den Öl-produzierenden Organismen gehören, das heißt die für die Herstellung von Ölen verwendet werden, wie Mikroorganismen wie *Cryptocodinium*, *Thraustochytrium*, *Phaeodactylum* und *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodinium* sowie andere Algen oder Pilze sowie Tiere oder Pflanzen, insbesondere Pflanzen bevorzugt Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten, wie Sojabohne, Erdnuss, Raps, Canola, Sonnenblume, Safflor, Nachtkerze, Lein, Soja, Borretsch, Bäume (Ölpalme, Kokosnuss) oder Feldfrüchte, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa oder Buschpflanzen (Kaffee, Kakao, Tee), *Salix*-Arten, Bäume (Ölpalme, Kokosnuß) sowie ausdauernde Gräser und Futterfeldfrüchte. Besonders bevorzugte erfindungsgemäße Pflanzen sind Ölfruchtpflanzen, wie Soja, Erdnuß, Raps, Canola, Lein, Nachtkerze, Sonnenblume, Safflor oder Bäume (Ölpalme, Kokosnuß).

30

Das Verfahren beinhaltet entweder die Züchtung eines geeigneten transgenen Organismus bzw. transgenen Mikroorganismus oder die Züchtung von transgenen Pflanzenzellen, -geweben, -organen oder ganzen Pflanzen, umfassend die erfindungsgemäßen Nukleotidsequenzen der SEQ ID NO: 1, 3, 5 oder 11 gegebenenfalls in Verbindungen mit den in SEQ ID NO: 7 und/oder SEQ ID NO: 9 dargestellten Sequenzen allein oder in Kombination mit Sequenzen von vorteilhaften erfindungsgemäßen Expressionskassetten in vorteilhaften Vektoren mit SEQ ID NO: 13-17 oder ihre Homologen, Derivate oder Analoga oder ein Genkonstrukt, das die SEQ ID NO: 1, 3, 5 oder 11 ggf. in Verbindung mit SEQ ID NO: 7 und/oder 9 oder ihre Homologen, Derivate oder Analoga umfasst, oder einen Vektor, der diese Sequenz oder das Genkonstrukt umfasst, welches die Expression erfindungsgemäßer Nukleinsäuremoleküle herbeiführt, so dass eine Feinchemikalie produziert wird. Bei einer bevorzugten Ausführungsform umfasst das Verfahren ferner den Schritt des Gewinnens einer Zelle, die eine solche erfindungsgemäße Nukleinsäuresequen-

## 10

zen enthält, wobei eine Zelle mit einer Desaturasenukleinsäuresequenz, einem Genkonstrukt oder einem Vektor, welche die Expression einer erfindungsgemäßen Desaturasenukleinsäure allein oder in Kombination herbeiführen, transformiert wird. Bei einer  
5 weiteren bevorzugten Ausführungsform umfasst dieses Verfahren ferner den Schritt des Gewinnens der Feinchemikalie aus der Kultur. Bei einer besonders bevorzugten Ausführungsform gehört die Zelle zur Ordnung der Ciliaten, zu Mikroorganismen, wie Pilzen, oder zum Pflanzenreich, insbesondere zu Ölfruchtpflanzen,  
10 besonders bevorzugt sind Mikroorganismen oder Ölfruchtpflanzen beispielsweise Erdnuss, Raps, Canola, Lein, Soja, Safflower (Distel), Sonnenblumen oder Borretsch.

Unter transgen im Sinne der Erfindung ist zu verstehen, daß die  
15 im Verfahren verwendeten Nukleinsäuren oder die erfindungsgemäßen Expressionskassetten nicht an ihrer natürlichen Stelle im Genom eines Organismus sind, dabei können die Nukleinsäuren homolog oder heterolog exprimiert werden. Transgen bedeutet aber auch, dass die Nukleinsäuren oder Expressionskassetten an ihrem natürlichen Platz im Genom eines Organismus sind, dass jedoch die Sequenz gegenüber der natürlichen Sequenz verändert wurde und/oder  
20 dass die Regulationssequenzen, der natürlichen Sequenzen verändert wurden. Bevorzugt ist unter transgen die Expression der erfindungsgemäßen Nukleinsäuren an nicht natürlicher Stelle im Genom  
25 zu verstehen, das heißt eine homologe oder bevorzugt heterologe Expression der Nukleinsäuren liegt vor. Bevorzugte transgene Organismen sind die oben genannten transgenen Pflanzen bevorzugt Ölfruchtpflanzen.

30 Aus den im Verfahren hergestellten Fettsäureestern lassen sich die enthaltenden mehrfach ungesättigten Fettsäuren beispielsweise über eine Alkalibehandlung wie wässrige KOH oder NaOH vorteilhaft in Gegenwart eines Alkohols wie Methanol oder Ethanol freisetzen und isolieren über beispielsweise Phasentrennung und anschließen-  
35 der Ansäuerung über z.B.  $H_2SO_4$ .

Die im Verfahren hergestellten Fettsäureester fallen in Form von Ölen, Lipiden und/oder Fettsäuren an, die mindestens zwei Doppelbindungen in den Fettsäuremolekülen bevorzugt drei, vier, fünf  
40 oder sechs Doppelbindungen enthalten. Auch sind Zusammensetzungen, die die genannten Öl-, Lipid- und/oder Fettsäuren enthalten, sowie die Verwendung der Öle, Lipide und/oder Fettsäuren oder der Zusammensetzungen in Futter, Nahrungsmitteln, Kosmetika oder Pharmazeutika ein weiterer Anwendungsmöglichkeit der vorgenannten  
45 Stoffe.

## 11

- Ein weiterer Aspekt betrifft ein Verfahren zur Modulation der Produktion eines Moleküls durch einen Mikroorganismus. Diese Verfahren umfassen das Zusammenbringen der Zelle mit einer Substanz, welche die erfindungsgemäßen Desaturaseaktivität allein oder in
- 5 Kombination oder die Desaturasenukleinsäureexpression moduliert, so dass eine zellassozierte Aktivität relativ zu der gleichen Aktivität in Abwesenheit der Substanz verändert wird. Bei einer bevorzugten Ausführungsform wird/werden ein oder zwei Stoffwechselweg(e) der Zelle für Lipide und Fettsäuren, Cofaktoren und En-
- 10 zyme moduliert oder der Transport von Verbindungen über diese Membranen moduliert, so dass die Ausbeute oder die Rate der Produktion einer gewünschten Feinchemikalie durch diesen Mikroorganismus verbessert ist. Die Substanz, welche die Desaturaseaktivität moduliert, kann eine Substanz sein, welche die
- 15 Desaturaseaktivität oder Desaturasenukleinsäureexpression stimuliert oder die als Zwischenprodukt bei der Fettsäurebiosynthese verwendet werden kann. Beispiele für Substanzen, welche die Desaturaseaktivität oder Desaturasenukleinsäureexpression stimulieren, sind u.a. kleine Moleküle, aktive Desaturasen sowie
- 20 desaturasenkodierende Nukleinsäuren, die in die Zelle eingebracht worden sind. Beispiele für Substanzen, welche die Desaturaseaktivität oder -Expression hemmen, sind u.a. kleine Moleküle und Antisense- Desaturasenukleinsäuremoleküle.
- 25 Ein weiterer Aspekt betrifft ein Verfahren zur Modulation der Ausbeuten einer gewünschten Verbindung aus einer Zelle, umfassend das Einbringen eines Wildtyp- oder Mutanten-Desaturasegens, das entweder auf einem separaten Plasmid gehalten oder in das Genom der Wirtszelle integriert wird, in eine Zelle. Bei Integration in
- 30 das Genom kann die Integration zufallsgemäß sein oder durch derartige Rekombination erfolgen, dass das native Gen durch die eingebrachte Kopie ersetzt wird, wodurch die Produktion der gewünschten Verbindung durch die Zelle moduliert wird, oder durch Verwendung eines Gens in trans, so dass das Gen mit einer funktionellen Expressionseinheit, welche mindestens eine die Expression eines Gens gewährleistende Sequenz und mindestens eine die Polyadenylierung eines funktionell transkribierten Gens gewähr-
- 35 leistende Sequenz enthält, funktionell verbunden ist.
- 40 Bei einer bevorzugten Form des Verfahrens sind die Ausbeuten modifiziert. Bei einer weiteren bevorzugten Ausführungsform wird die gewünschte Chemikalie vermehrt, wobei unerwünschte störende Verbindungen vermindert werden können. Bei einer besonders bevorzugten Ausführungsform ist die gewünschte Fein-
- 45 chemikalie ein Lipid oder eine Fettsäure, ein Cofaktor oder ein Enzym. Bei besonders bevorzugten Ausführungsform ist diese Chemikalie eine mehrfach ungesättigte Fettsäure. Stärker bevor-

## 12

zugt ist sie ausgewählt aus Arachidonsäure (ARA), Eicosapentaensäure (EPA) oder Docosahexaensäure (DHA).

Die vorliegende Erfindung stellt vorteilhafte Multiexpressions-  
5 kassetten und Konstrukte zur multiparallelen samenspezifischen Expression von Genkombinationen in Pflanzen zur Verfügung.

Diese können im oben beschriebenen Verfahren zur Expression von Genen, bevorzugt den in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5,  
10 SEQ ID NO: 7, SEQ ID NO: 9 oder SEQ ID NO: 11 beschriebenen, in Algen und Pilzen und Pflanzen, insbesondere Ölfruchtpflanzen sind bevorzugte Organismen für das Verfahren verwendet werden.

Durch die Verwendung der erfindungsgemäßen Expressionskassetten  
15 können im Verfahren in Verbindung mit den oben genannten Nukleinsäuremoleküle zur gentechnologisch Veränderung von Pflanzen verwendet werden, so dass sie schließlich zur Herstellung von besseren oder effizienteren Produzenten einer oder mehrerer Feinchemikalien führen. Diese verbesserte Produktion oder Effizienz der  
20 Produktion einer Feinchemikalie kann durch eine direkte Wirkung der Manipulation eines erfindungsgemäßen Gens oder durch eine indirekte Wirkung dieser Manipulation hervorgerufen werden. Unter Feinchemikalien sind im Sinne der Erfindung beispielsweise Fettsäureester, die mehrfach ungesättigte Fettsäuren mit mindestens  
25 zwei Doppelbindungen enthalten wie Sphingolipide, Phosphoglyceride, Lipide, Glycolipide, Phospholipide, Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Fettsäureester, die die mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen enthalten, zu verstehen. Weiter sind darunter zu ver-  
30 stehen Verbindungen wie Vitamine beispielsweise Vitamin E, Vitamin C, Vitamin B2, Vitamin B6, Pantolacton, Carotinoide wie Astaxanthin,  $\beta$ -Carotin, Zeaxanthin und andere.

Moose und Algen sind die einzigen bekannten Pflanzensysteme,  
35 die erhebliche Mengen an mehrfach ungesättigten Fettsäuren, wie Arachidonsäure (ARA) und/oder Eicosapentaensäure (EPA) und/oder Docosahexaensäure (DHA) herstellen. Moose enthalten PUFAs in Membranlipiden während Algen, algenverwandte Organismen und einige Pilze auch nennenswerte Mengen an PUFAs in der Triacyl-  
40 glycerolfraktion akkumulieren. Daher eignen sich Nukleinsäuremoleküle, die aus solchen Stämmen isoliert werden, die PUFAs auch in der Triacylglycerolfraktion akkumulieren, besonders vorteilhaft zur Modifikation des Lipid- und PUFA-Produktionssystems in einem Wirt, insbesondere in Mikroorganismen, wie den vorstehend  
45 erwähnten Mikroorganismen, und Pflanzen, wie Ölfruchtpflanzen,

## 13

beispielsweise Raps, Canola, Lein, Soja, Sonnenblumen, Borretsch. Sie sind deshalb vorteilhaft im Verfahren verwendbar.

Die im Verfahren unter Verwendung der erfindungsgemäßen Expressionskassetten verwendeten Nukleinsäuresequenzen codieren für Desaturasen, die zur Produktion langkettiger mehrfach ungesättigter Fettsäuren, vorzugsweise mit mehr als sechzehn, achtzehn oder zwanzig Kohlenstoffatomen im Kohlenstoffgrundgerüst der Fettsäure und/oder mindestens zwei Doppelbindungen in der Kohlenstoffkette, geeignet sind, wobei eine Nukleinsäure für ein Enzym codiert, das Doppelbindungen in die  $\Delta$ -5-Position, in einem anderen Fall in die  $\Delta$ -6-Position und in einem weiteren Fall in die  $\Delta$ -12-Position einführen kann. Mithilfe dieser Nukleinsäuren können hohe Mengen an PUFAs in der Triacylglycerolfraktion erhalten werden. Weiterhin wurden weitere Desaturasen isoliert, die allein oder zusammen mit einer  $\Delta$ -4-Desaturase für ein Verfahren zur Produktion polyungesättigter Fettsäuren genutzt werden können. Dabei ist in der Anmeldung unter dem Singular d.h. unter einem Desaturasegen oder -Protein auch der Plural d.h. die Desaturasegenen oder -Proteinen zu verstehen.

Die Herstellung einer Triensäure mit  $C_{18}$ -Kohlenstoffkette mithilfe von Desaturasen konnte bisher gezeigt werden. In diesen literaturbekannten Verfahren wurde die Herstellung von  $\gamma$ -Linolensäure beansprucht. Bisher konnte jedoch niemand die Herstellung sehr langkettiger mehrfach ungesättigter Fettsäuren (mit  $C_{20}$ - und längerer Kohlenstoffkette sowie von Triensäuren und höher ungesättigten Typen) allein durch modifizierte Organismen zeigen.

Zur Herstellung der erfindungsgemäßen langkettiger PUFAs müssen die mehrfach ungesättigten  $C_{18}$ -Fettsäuren zunächst durch die enzymatische Aktivität einer Elongase um mindestens zwei Kohlenstoffatome verlängert werden. Nach einer Elongationsrunde führt diese Enzymaktivität zu  $C_{20}$ -Fettsäuren, und nach zwei, drei und vier Elongationsrunden zu  $C_{22}$ -,  $C_{24}$ - oder  $C_{26}$ -Fettsäuren. Die in dieser Erfindung offenbarten Nukleinsäuresequenzen, die für verschiedene Desaturasen codieren, können im Konzert mit Elongasen zu sehr langkettigen, polyungesättigten führen. Die Aktivität der erfindungsgemäßen Desaturasen führt vorzugsweise zu  $C_{18}$ -,  $C_{20}$ - und/oder  $C_{22}$ -Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier, fünf oder sechs Doppelbindungen, besonders bevorzugt zu  $C_{18}$ - und/oder  $C_{20}$ -Fettsäuren mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise mit drei, vier oder fünf Doppelbindungen im Molekül. Die Fettsäureelongation kann durch Kombination der erfindungsgemäßen Desaturasen mit einer Elongaseaktivität erfolgen, wobei die durch die in SEQ ID NO: 9 codierte Elongase vorteilhaft verwendet

## 14

- werden kann. Nachdem die Verlängerung mit dem erfindungsgemäßen Enzym(en) stattgefunden hat, können weitere Desaturierungsschritte wie z.B. eine solche in  $\Delta$ -5-Position erfolgen. Auch die Kombination mit anderen Elongasen wie solche, die zu einer Ver-  
5 längerung von  $C_{18}$ - auf  $C_{20}$ - oder von  $C_{20}$ - auf  $C_{22-24}$  Ketten wie in WO0012720 offenbart führt, kann Verwendung finden und/oder einer Desaturase mit Aktivität für  $\Delta$ -4-Position kann vorteilhaft eingesetzt werden, um die hoch desaturierten Fettsäuren zu erhalten. Daher führen die Produkte der Desaturaseaktivitäten und der  
10 möglichen weiteren Desaturierung zu bevorzugten PUFAs mit einem höheren Desaturierungsgrad, wie Dihomo-gamma-Linolensäure, Docosadiensäure, Arachidonsäure,  $\omega$ 6-Eicosatriendihomo- $\gamma$ -linolensäure, Eicosapentaensäure,  $\omega$ 3-Eicosatriensäure,  $\omega$ 3-Eicosatetraensäure, Docosapentaensäure oder Docosahexaensäure. Substrate der  
15 erfindungsgemäßen Enzymaktivität sind zum Beispiel Taxolsäure; 6,9-Octadecadiensäure, Linolsäure, Pinolensäure,  $\alpha$ -oder  $\gamma$ -Linolensäure oder Stearidonsäure sowie Arachidonsäure, Eicosatetraensäure, Docosapentaensäure, Eicosapentaensäure. Bevorzugte Substrate sind Linolsäure,  $\gamma$ -Linolensäure und/oder  $\alpha$ -Linolensäure  
20 sowie Arachidonsäure, Eicosatetraensäure, Docosapentaensäure, Eicosapentaensäure. Besonders bevorzugt als Produkte des erfindungsgemäßen Verfahrens sind Arachidonsäure, Docosapentaensäure, Eicosapentaensäure. Die  $C_{18}$ -Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungs-  
25 gemäße enzymatische Aktivität in Form der freien Fettsäure oder in Form der Ester, wie Phospholipide, Glycolipide, Sphingolipide, Phosphoglyceride, Monoacylglyceride, Diacylglyceride oder Triacylglyceride, verlängert werden.
- 30 Für die menschliche Ernährung ist konjugierte Linolsäure "CLA" von besonderer Bedeutung. Unter CLA versteht man insbesondere Fettsäuren wie  $C_{18:2}$  9 cis, 11trans oder das Isomer  $C_{18:2}$  10trans, 12 cis, die aufgrund menschlicher Enzymsysteme nach Aufnahme im Körper desaturiert bzw. elongiert werden können und zu gesundheits-  
35 fördernden Effekten beitragen können. Mit den erfindungsgemäßen Desaturasen ( $\Delta$ -12-Desaturase) können auch solche konjugierten Fettsäuren mit wenigstens zwei Doppelbindungen im Molekül desaturiert werden und damit solche gesundheitsfördernden Fettsäuren der menschlichen Ernährung zugeführt werden. Weitere  
40 Beispiel für konjugierte Fettsäuren sind alpha-Parinarinsäure, Punicasäure, Eleostearinsäure und Calendulasäure.

Unter der Verwendung von der erfindungsgemäßen Expressionkassetten in Klonierungsvektoren in Pflanzen und bei der Pflanzentrans-  
45 formation, wie denjenigen, die veröffentlicht sind in und dort zitiert sind: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapitel 6/7, S. 71-119 (1993); F.F.

## 15

White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225)), lassen sich Gene vorteilhaft Biosynthesegene wie die oben beschriebenen Nukleinsäuren zur gentechnologischen Veränderung eines breiten Spektrums an Pflanzen verwenden, so dass diese ein besserer oder effizienterer Produzent beispielsweise eines oder mehrerer von Lipiden hergeleiteter Produkte, wie PUFAs, werden. Diese verbesserte Produktion oder Effizienz der Produktion eines beispielsweise von Lipiden hergeleiteten Produktes, wie PUFAs, kann durch direkte Wirkung der Manipulation oder eine indirekte Wirkung dieser Manipulation hervorgerufen werden.

Es gibt eine Reihe von Mechanismen, durch die die Veränderung eines erfindungsgemäßen Desaturaseproteins die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie aus einer Ölfruchtpflanze oder einem Mikroorganismus aufgrund eines veränderten Proteins direkt beeinflussen kann. Die Anzahl oder Aktivität des Desaturaseproteins oder -Gens sowie von Genkombinationen von Desaturasen und Elongasen kann erhöht sein, so dass größere Mengen dieser Verbindungen de novo hergestellt werden, weil den Organismen diese Aktivität und Fähigkeit zur Biosynthese vor dem Einbringen des entsprechenden Gens fehlte. Entsprechendes gilt für die Kombination mit weiteren Desaturasen oder Elongasen oder weiteren Enzymen aus dem Lipidstoffwechsel. Auch die Verwendung verschiedener divergenter, d.h. auf DNA-Sequenzebene unterschiedlicher Sequenzen kann dabei vorteilhaft sein bzw. die Verwendung von Promotoren zur Genexpression, die eine andere zeitliche Genexpression z.B. abhängig vom Reifegrad eines Samens oder Öl-speichernden Gewebes ermöglicht.

Durch das Einbringen eines Desaturasegens oder mehrerer Desaturasegene unter Kontrolle der erfindungsgemäßen Expressionskassetten in einen Organismus allein oder in Kombination mit anderen Genen in eine Zelle kann nicht nur den Biosynthesefluss zum Endprodukt erhöht, sondern auch die entsprechende Zusammensetzung der Enprodukte beispielsweise der Triacylglycerine erhöht oder de novo geschaffen werden. Ebenso kann die Anzahl oder Aktivität anderer Gene, die am Import von Nährstoffen, die zur Biosynthese einer oder mehrerer Feinchemikalien (z.B. Fettsäuren, polaren und neutralen Lipiden) nötig sind, erhöht sein, so dass die Konzentration dieser Vorläufer, Cofaktoren oder Zwischenverbindungen innerhalb der Zellen oder innerhalb des Speicherkompartiments er-

hört ist, wodurch die Fähigkeit der Zellen zur Produktion von PUFAs, wie im folgenden beschrieben, weiter gesteigert wird. Fettsäuren und Lipide sind selbst als Feinchemikalien wünschenswert; durch Optimierung der Aktivität oder Erhöhung der Anzahl einer  
5 oder mehrerer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Zerstören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen  
10 aus Pflanzen oder Mikroorganismen zu steigern.

Die Mutagenese der/des erfindungsgemäßen Desaturasegene(s) kann weiterhin zu einem Desaturaseprotein mit geänderten Aktivitäten führen, welche die Produktion einer oder mehrerer gewünschter  
15 Feinchemikalien direkt oder indirekt beeinflussen. Beispielsweise kann die Anzahl oder Aktivität der/des erfindungsgemäßen Desaturasegens(e) gesteigert werden, so dass die normalen Stoffwechselabfälle oder -nebenprodukte der Zelle (deren Menge möglicherweise aufgrund der Überproduktion der gewünschten Feinchemikalie erhöht ist) effizient exportiert werden, bevor sie  
20 andere Moleküle oder Prozesse innerhalb der Zelle (welche die Lebensfähigkeit der Zelle senken würden) zerstören oder die Biosynthesewege der Feinchemikalie stören würden (wodurch die Ausbeute, Produktion oder Effizienz der Produktion der gewünschten  
25 Feinchemikalie verringert wird). Ferner können die relativ großen intrazellulären Mengen der gewünschten Feinchemikalie selbst toxisch für die Zelle sein oder Enzym-Rückkopplungsmechanismen, wie die allosterische Regulation, stören, beispielsweise könnte sie durch Steigerung der Aktivität oder Anzahl anderer strom-  
30 abwärts folgender Enzyme oder Entgiftungsenzyme des PUFA-Wegs die Allokation der PUFA in die Triacylglycerin-Fraktion steigern, man könnte die Lebensfähigkeit von Saatzellen erhöhen, was wiederum zu besserer Entwicklung von Zellen in Kultur oder zu Saaten führt, die die gewünschte Feinchemikalie produzieren.  
35 Das erfindungsgemäße Desaturasegen kann auch so manipuliert werden, dass die entsprechenden Mengen der verschiedenen Lipid- und Fettsäuremoleküle hergestellt werden. Dies kann eine einschneidende Wirkung auf die Lipidzusammensetzung der Membran der Zelle haben und erzeugt neue Öle zusätzlich zum Auftreten  
40 neusynthetisierter PUFAs. Da jeder Lipidtyp unterschiedliche physikalische Eigenschaften hat, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität erheblich verändern. Änderungen der Membranfluidität können sich auf den Transport von Molekülen über die Membran sowie auf die  
45 Unversehrtheit der Zelle auswirken, die beide eine entscheidende Wirkung auf die Produktion von Feinchemikalien besitzen. In Pflanzen können diese Änderungen überdies auch andere Merk-

male, wie Toleranz gegenüber abiotischen und biotischen Stress-situationen, beeinflussen.

Im Verfahren können isolierte Nukleinsäuremoleküle (z.B. cDNAs),  
5 umfassend Nukleotidsequenzen, die eine Desaturase oder mehrere  
Desaturasen oder biologisch aktive Teile davon codieren, oder Nu-  
kleinsäurefragmente, die sich als Primer oder Hybridisierungs-  
sonden zum Nachweis oder zur Amplifikation desaturasekodierender  
Nukleinsäuren (z.B. DNA oder mRNA) eignen, verwendet werden. Bei  
10 besonders bevorzugten Ausführungsformen umfasst das Nukleinsäure-  
molekül eine der in Sequenz ID NO:1 bzw 3 und 5 dargestellten Nu-  
kleotidsequenzen oder die kodierende Region oder ein Komplement  
einer dieser Nukleotidsequenzen. Bei anderen besonders bevorzug-  
ten Ausführungsformen umfasst das isolierte Nukleinsäuremolekül  
15 eine Nukleotidsequenz, die an eine Nukleotidsequenz, wie in der  
Sequenz SEQ ID NO: 1, 3, 5 oder 11 dargestellt, oder einen Teil  
davon hybridisiert oder zu mindestens etwa 50 %, vorzugsweise  
mindestens etwa 60 %, stärker bevorzugt mindestens etwa 70 %, 80 %  
oder 90 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr homolog dazu ist. Bei anderen  
20 bevorzugten Ausführungsformen kodiert das isolierte Nukleinsäure-  
molekül eine der in der Sequenz SEQ ID NO: 2, 4, 6 oder 12 darge-  
stellten Aminosäuresequenzen. Das bevorzugte Desaturasegen be-  
sitzt vorzugsweise auch mindestens eine der hier beschriebenen  
25 Desaturaseaktivitäten.

Bei einer weiteren Ausführungsform kodiert das isolierte  
Nukleinsäuremolekül ein Protein oder einen Teil davon, wobei  
das Protein oder der Teil davon eine Aminosäuresequenz enthält,  
30 die ausreichend homolog zu einer Aminosäuresequenz der Sequenz  
SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil  
davon eine Desaturaseaktivität beibehält. Vorzugsweise behält das  
Protein oder der Teil davon, das/der von dem Nukleinsäuremolekül  
kodiert wird, die Fähigkeit, am Stoffwechsel von zum Aufbau  
35 von Zellmembranen von Pflanzen notwendigen Verbindungen oder  
am Transport von Molekülen über diese Membranen teilzunehmen.  
Bei einer Ausführungsform ist das von dem Nukleinsäuremolekül  
kodierte Protein zu mindestens etwa 50 %, vorzugsweise mindestens  
etwa 60 % und stärker bevorzugt mindestens etwa 70 %, 80 % oder  
40 90 % und am stärksten bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr homolog zu einer Aminosäuresequenz der  
Sequenz SEQ ID NO: 2, 4, 6 oder 12. Bei einer weiteren bevor-  
zugten Ausführungsform ist das Protein ein Voll-längen-Protein,  
das im wesentlichen in Teilen homolog zu einer gesamten Amino-  
45 säuresequenz der SEQ ID NO: 2, 4, 6 oder 12 (die von dem in

SEQ ID NO: 1, 3, 5 oder 11 gezeigten offenen Leserahmen herrührt) ist.

Bei anderen Ausführungsformen umfasst die isolierte Desaturase  
5 eine Aminosäuresequenz, die zu mindestens etwa 50 % homolog zu einer der Aminosäuresequenzen der SEQ ID NO: 2, 4, 6 oder 12 ist und am Stoffwechsel von zum Aufbau von Fettsäuren in einem Mikroorganismus oder einer Pflanzenzelle notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann,  
10 wobei desaturierte C<sub>18</sub>- oder C<sub>20-22</sub>-Kohlenstoffketten mit Doppelbindungen an mindestens zwei Stellen gemeint ist.

Bei einer anderen bevorzugten Ausführungsform rührt das isolierte Nukleinsäuremolekül von *Phaeodactylum tricornutum* UTEX646 her  
15 und kodiert ein Protein (z.B. ein Desaturasefusionsprotein), das eine biologisch aktive Domäne enthält, die zu mindestens etwa 50 % oder mehr homolog zu einer Aminosäuresequenz der Sequenz SEQ ID NR 2, 4, 6 oder 12 ist und die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen von Pflanzen notwendigen Ver-  
20 bindungen oder am Transport von Molekülen über diese Membranen teilzunehmen, beibehält oder zumindest eine der Desaturierungsaktivitäten resultierend in PUFAs wie GLA, ALA, Dihomo-gamma Linolensäure, ARA, EPA oder DHA oder deren Vorläufermoleküle besitzt, und umfasst auch heterologe Nukleinsäuresequenzen, die  
25 ein heterologes Polypeptid oder regulatorische Proteine kodieren.

Alternativ kann die isolierte Desaturase eine Aminosäuresequenz umfassen, die von einer Nukleotidsequenz kodiert wird, die an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert, z.B. unter stringenten Bedingungen hybridisiert, oder zu  
30 mindestens etwa 50 %, vorzugsweise mindestens etwa 60 %, stärker bevorzugt mindestens etwa 70 %, 80 % oder 90 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder mehr homolog dazu ist. Es ist ebenfalls bevorzugt, dass die bevorzugten Desaturaseformen ebenfalls eine der hier beschriebenen  
35 Desaturaseaktivitäten besitzen.

Bei einer anderen Ausführungsform ist das isolierte Nukleinsäuremolekül mindestens 15, 25, 50, 100, 250 oder mehr Nukleotide lang  
40 und hybridisiert unter stringenten Bedingungen an ein Nukleinsäuremolekül, das eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 17 umfasst. Vorzugsweise entspricht das isolierte Nukleinsäuremolekül einem natürlich vorkommenden Nukleinsäuremolekül. Stärker bevorzugt kodiert das isolierte Nukleinsäuremolekül  
45 natürlich vorkommende *Phaeodactylum*-Desaturase oder einen biologisch aktiven Teil davon.

Eine weitere Ausführungsform der Erfindung sind Expressionskassetten, die die Expression der erfindungsgemäßen Nukleinsäuren mit den Sequenzen SEQ ID NO: 1, 3, 5 oder 11 in den verschiedenen Organismen wie pflanzliche Zellen, Geweben, Teilen von Pflanzen 5 oder ganzen Pflanzen ermöglichen.

Unter der erfindungsgemäßen Expressionskassette (= Nukleinsäurekonstrukt oder -fragment) sind die in SEQ ID NO: 32 als L1 und einem Promotor, einem Strukturgen ausgewählt aus den vorteilhaften Sequenzen SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder 10 SEQ ID NO: 11, die sich als Ergebnis des genetischen Codes und/oder deren funktionellen oder nicht funktionellen Derivate zu verstehen, und die Polylinker-Terminator-Polylinker-Sequenzen (= L2) SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35 zu verstehen. 15 Diese steuern vorteilhafterweise die Genexpression in der Wirtszelle. Diese in den Konstrukten enthaltenen regulatorischen Sequenzen sollen die gezielte Expression der Gene und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtorganismus bedeuten, dass das Gen erst nach Induktion exprimiert 20 und/oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird. Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Das Genkonstrukt kann aber auch einfacher aufgebaut 30 sein, das heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz oder dessen Derivate inseriert und der natürliche Promotor mit seiner Regulation wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und/oder die Genexpression gesteigert wird. Diese veränderten Promotoren können in Form 35 von Teilsequenzen (= Promotor mit Teilen der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen zur Steigerung der Aktivität gebracht werden. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte 40 "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die  $\Delta$ -5-Desaturase-/ $\Delta$ -6-Desaturase 45 und/oder  $\Delta$ -12-Desaturasegene können in einer oder mehreren

Kopien in der Expressionskassette (= Genkonstrukt) enthalten sein.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie  
5 oben beschrieben vorzugsweise die Genexpression der eingeführten  
Gene positiv beeinflussen und dadurch erhöhen. So kann eine Ver-  
stärkung der regulatorischen Elemente vorteilhafterweise auf der  
Transkriptionsebene erfolgen, indem starke Transkriptionssignale  
wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist  
10 aber auch eine Verstärkung der Translation möglich, indem bei-  
spielsweise die Stabilität der mRNA verbessert wird.

Ein weiterer Aspekt der Erfindung betrifft Vektoren, z.B.  
rekombinante Expressionsvektoren, die mindestens eine der erfin-  
15 dungsgemäßen Expressionskassetten enthalten, und Wirtszellen, in  
die die erfindungsgemäßen Expressionskassetten oder diese Vekto-  
ren eingebracht worden sind, insbesondere Mikroorganismen, Pflan-  
zenzellen, Pflanzengewebe, -organe oder ganze Pflanzen. Bei einer  
Ausführungsform kann eine solche Wirtszelle Feinchemikalien-Ver-  
20 bindungen, insbesondere PUFAs, speichern; zur Isolation der ge-  
wünschten Verbindung werden die Zellen geerntet. Die Verbindung  
(Öle, Lipide, Triacylglyceride, Fettsäuren) oder die Desaturase  
können dann aus dem Medium oder der Wirtszelle, welche bei Pflan-  
zen Zellen sind, die Feinchemikalien enthalten oder speichern, am  
25 stärksten bevorzugt Zellen von Speichergeweben, wie Samenhüllen,  
Knollen, Epidermis- und Samenzellen, Endosperm oder Embryogewebe  
isoliert werden.

Noch ein weiterer Aspekt der Erfindung betrifft eine genetisch  
30 veränderte transgene Pflanze, bevorzugt ein Ölfruchtpflanze, wie  
vorstehend erwähnt, besonders bevorzugt eine Raps- oder Lein-  
pflanze, in die eine erfindungsgemäße Expressionskassette, die  
vorteilhaft weitere Gene wie Desaturasegen enthält, eingebracht  
worden ist. Bei einer Ausführungsform ist das Genom von Raps oder  
35 Lein durch Einbringen einer erfindungsgemäßen Expressionskassette  
vorteilhaft enthaltend weitere Nukleinsäuremoleküle, die bei-  
spielsweise eine Wildtyp- oder mutierte Desaturasesequenz kodiert,  
als Transgen verändert worden. Bei einer bevorzugten Ausführungs-  
form wird Raps oder Lein auch zur Produktion einer gewünschten  
40 Verbindung, wie Lipiden und Fettsäuren, wobei PUFAs besonders be-  
vorzugt sind, verwendet.

Bei noch einer weiteren bevorzugten Ausführungsform kann das  
Moos *Physcomitrella patens* zur Demonstration der Funktion einer  
45 Expressionskassette mit einem Desaturasegen unter Verwendung

## 21

homologer Rekombination auf der Basis der in dieser Erfindung beschriebenen Nukleinsäuren verwendet werden.

- Das Desaturasepolypeptid oder ein biologisch aktiver Teil davon  
5 kann vorteilhaft unter Kontrolle der erfindungsgemäßen Expressionskassette funktionsfähig mit einem weiteren Polypeptid, das eine andere enzymatische Aktivität als die Desaturasen hat beispielsweise eine Elongase-, Acyltransferase- oder sonstige Aktivität verbunden werden, so dass ein Fusionsprotein gebildet wird.  
10 Vorteilhaft hat dieses Fusionsprotein eine Aktivität, die sich von derjenigen der Desaturase allein unterscheidet. Bei anderen bevorzugten Ausführungsform nimmt dieses Fusionsprotein am Stoffwechsel von Verbindungen, die zur Synthese von Lipiden und Fettsäuren, Cofaktoren und Enzymen in Mikroorganismen oder  
15 Pflanzen notwendig sind, oder am Transport von Molekülen über diese Membranen teil. Besonders vorteilhaft moduliert das Einbringen dieses Fusionsproteins in einer Wirtszelle die Produktion einer gewünschten Verbindung innerhalb einer und durch die Zelle. Bei einer bevorzugten Ausführungsform enthalten diese Fusionsproteine auch  $\Delta$ -4-,  $\Delta$ -5- oder  $\Delta$ -6,  $\Delta$ -8-,  $\Delta$ -15,  $\Delta$ -17 oder  $\Delta$ -19-Desaturaseaktivitäten allein oder in Kombination. Insbesondere solche Genkombinationen sind bevorzugte Ausführungsformen, die aus SEQ ID NO: 7 oder 9 gewählt sind, bzw Teilen davon, Derivate oder ihren Homologen. Insbesondere solche  
20 Kombinationen sind bevorzugt, die die vollständige Proteinaktivität wie in SEQ ID NO: 1, 3, 5 oder 11 enthalten und in Multiexpressionskassetten definiert durch SEQ ID NO: 13, 14, 15, 16 und 17 eingefügt zur Transformation von Pflanzen und Expression in Pflanzen geeignet sind.

30

#### Eingehende Beschreibung der Erfindung

- Ein erfindungsgemäßer Gegenstand sind auch die Expressionskassetten in Verbindung mit isolierten Nukleinsäuresequenz(en), die für  
35 ein Polypeptid mit Desaturaseaktivität codiert, ausgewählt aus der Gruppe:

- a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,  
40
- b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten  
45 Aminosäuresequenzen erhalten werden,

## 22

- c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

Weiterhin betrifft die Erfindung eine Aminosäuresequenz, die durch die oben genannte(n) Nukleinsäuresequenz(en) codiert werden (für die Erfindung soll der Singular den Plural und umgekehrt umfassen). Speziell betrifft die Erfindung Aminosäuresequenzen, die durch die in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellte Sequenz codiert werden.

15

- Die vorliegende Erfindung stellt Expressionskassetten bereit, die für die Expression von Nukleinsäuren und Proteinmoleküle mit Desaturaseaktivität geeignet sind, und von Nukleinsäuren kodierend für Proteine, die am Stoffwechsel von Lipiden und Fettsäuren, PUFA-Cofaktoren und Enzymen in dem Moos *Physcomitrella patens* oder am Transport lipophiler Verbindungen über Membranen beteiligt sind. Die erfindungsgemäßen Verbindungen lassen sich zur Modulation der Produktion von Feinchemikalien aus Organismen wie Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis, Gerste, Sojabohne, Erdnuss, Baumwolle, Linum Arten wie Öl- oder Faserlein, Brassica-Arten, wie Raps, Canola und Rübsen, Pfeffer, Sonnenblume, Borretsch, Nachtkerze und Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Maniok, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölpalme, Kokosnuss) und ausdauernden Gräsern und Futterfeldfrüchten, entweder direkt (z.B. wenn die Überexpression oder Optimierung eines Fettsäurebiosynthese-Proteins einen direkten Einfluss auf die Ausbeute, Produktion und/oder Effizienz der Produktion der Fettsäure aus modifizierten Organismen hat) verwenden oder können eine indirekt Auswirkung haben, die dennoch zu einer Steigerung der Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Verbindung oder einer Abnahme unerwünschter Verbindungen führt (z.B. wenn die Modulation des Stoffwechsels von Lipiden und Fettsäuren, Cofaktoren und Enzymen zu Veränderungen der Ausbeute, Produktion und/oder Effizienz der Produktion oder der Zusammensetzung der gewünschten Verbindungen innerhalb der Zellen führt, was wiederum die Produktion einer oder mehrerer Feinchemikalien beeinflussen kann). Aspekte der Erfindung sind nachstehend weiter erläutert.

45

# I. Feinchemikalien und PUFAs

## 23

- Der Begriff "Feinchemikalie" ist im Fachgebiet bekannt und umfasst Moleküle, die durch einen Organismus produziert worden sind und Anwendungen in verschiedenen Industrien finden, wie, aber nicht beschränkt auf, die pharmazeutische, Landwirtschafts-,
- 5 Nahrungsmittel- und Kosmetik-Industrie. Diese Verbindungen umfassen Lipide, Fettsäuren, Cofaktoren und Enzyme usw. (wie z.B. beschrieben in Kuninaka, A. (1996) Nucleotides and related compounds, S. 561-612, in Biotechnology Bd. 6, Rehm et al., Hrsgb., VCH: Weinheim und darin enthaltenen Literaturstellen), Lipide,
- 10 gesättigte und ungesättigte Fettsäuren (z.B. Arachidonsäure), Vitamine und Cofaktoren (wie beschrieben in Ullmann's Encyclopedia of Industrial Chemistry, Bd. A27, Vitamins, S. 443-613 (1996) VCH: Weinheim und darin enthaltenen Literaturstellen; und Ong, A.S., Niki, E., & Packer, L. (1995) Nutrition, Lipids,
- 15 Health and Disease Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia and the Society for Free Radical Research - Asien, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCs Press (1995)), Enzyme und sämtliche anderen von Gutcho (1983) in Chemicals by Fermentation,
- 20 Noyes Data Corporation, ISBN: 0818805086, und darin angegebenen Literaturstellen beschriebenen Chemikalien. Der Stoffwechsel und die Verwendungen bestimmter Feinchemikalien sind nachstehend weiter erläutert.
- 25 Die Kombination verschiedener Vorläufermoleküle und Biosyntheseenzyme führt zur Herstellung verschiedener Fettsäuremoleküle, was eine entscheidende Auswirkung auf die Zusammensetzung der Membran hat. Es kann angenommen werden, dass PUFAs nicht nur einfach in Triacylglycerin, sondern auch in Membranlipide eingebaut werden.
- 30 Vorläufer für die PUFA-Biosynthese sind beispielsweise Ölsäure, Linol- und Linolensäure. Diese C<sub>18</sub>-Kohlenstoff-Fettsäuren müssen auf C<sub>20</sub> und C<sub>22</sub> verlängert werden, damit Fettsäuren vom Eicosa- und Docosa-Kettentyp erhalten werden. Mithilfe
- 35 verschiedener Desaturasen, wie Enzymen, welche  $\Delta$ -12-Desaturase,  $\Delta$ -15-Desaturase,  $\Delta$ -6-Desaturase-,  $\Delta$ -5- und  $\Delta$ -4-Desaturaseaktivität aufweisen, können Arachidonsäure, Eicosapentaensäure und Docosahexaensäure sowie verschiedene andere langkettige PUFAs erhalten, extrahiert und für verschiedene Zwecke bei Nahrungs-
- 40 mittel-, Futter-, Kosmetik- oder pharmazeutischen Anwendungen verwendet werden.
- Zur Herstellung langkettiger PUFAs müssen, wie oben erwähnt, die mehrfach ungesättigten C<sub>18</sub>- bzw C<sub>20</sub>-Fettsäuren mehrfach
- 45 desaturiert werden. Die erfindungsgemäßen Nukleinsäuresequenzen kodieren erste funktionell aktive Desaturasen aus *Phaeodactylum tricornerum*, einem Mikroorganismus, der PUFAs in der Triacyl-

- glycerolfraktion enthält. Mit den erfindungsgemäßen Desaturasen können Doppelbindungen in die  $\Delta$ -5-,  $\Delta$ -6- oder  $\Delta$ -12-Position eingeführt werden. Die Aktivitäten der erfindungsgemäßen Desaturasen führt vorzugsweise zu  $C_{18}$ - +  $C_{20}$ -Fettsäuren mit mindestens zwei, 5 drei, vier oder fünf Doppelbindungen im Fettsäuremolekül, vorzugsweise zu  $C_{20}$ -Fettsäuren mit vorteilhaft drei, vier oder fünf Doppelbindungen im Fettsäuremolekül. Die Desaturierung kann vor oder nach Elongation der entsprechenden Fettsäure erfolgen. Daher führen die Produkte der Desaturaseaktivitäten und der möglichen weiteren Desaturierung und Elongation zu bevorzugten PUFAs mit höherem Desaturierungsgrad, einschließlich einer weiteren Elongation von  $C_{20}$  zu  $C_{22}$ -Fettsäuren, zu Fettsäuren wie Linolsäure, Docosadiensäure, dihomog-linolensäure, Arachidonsäure,  $\omega$ 6-Eicosatriendihomog-linolensäure, Eicosapentaensäure,  $\omega$ 3-Eicosatriensäure, 15  $\omega$ 3-Eicosatetraensäure, Docosapentaensäure oder Docosahexaensäure. Substrate dieser erfindungsgemäßen Enzymaktivität sind zum Beispiel Taxolsäure, 6,9-Octadecadiensäure, Ölsäure, Linolsäure,  $\gamma$ -Linolensäure, Pinolensäure,  $\alpha$ -Linolensäure, Arachidonsäure, Eicosapentaensäure oder Stearidonsäure. Bevorzugte 20 Substrate sind Linolsäure,  $\gamma$ -Linolensäure und/oder  $\alpha$ -Linolensäure, dihomog-linolensäure bzw. Arachidonsäure, Eicosatetraensäure oder Eicosapentaensäure. Die  $C_{18}$ -oder  $C_{20}$ -Fettsäuren mit mindestens zwei Doppelbindungen in der Fettsäure können durch die erfindungsgemäße Enzymaktivität in Form der freien Fettsäure oder 25 in Form der Ester, wie Phospholipide, Glykolipide, Sphingolipide, Phosphoglyceride, Monoacylglyceride, Diacylglyceride, Triacylglyceride oder sonstige Ester, verlängert werden.
- Ferner müssen Fettsäuren anschließend an verschiedene Modifikationsorte transportiert und in das Triacylglycerin- 30 Speicherlipid eingebaut werden. Ein weiterer wichtiger Schritt bei der Lipidsynthese ist der Transfer von Fettsäuren auf die polaren Kopfgruppen, beispielsweise durch Glycerin-Fettsäure-Acyltransferase (siehe Frentzen, 1998, Lipid, 100(4-5):161-166).
- 35 Veröffentlichungen über die Pflanzen-Fettsäurebiosynthese, Desaturierung, den Lipidstoffwechsel und Membrantransport von fetthaltigen Verbindungen, die Betaoxidation, Fettsäuremodifikation und Cofaktoren, Triacylglycerin-Speicherung und -Assemblierung einschließlich der Literaturstellen darin siehe 40 in den folgenden Artikeln: Kinney, 1997, Genetic Engineering, Hrsgb.: JK Setlow, 19:149-166; Ohlrogge und Browse, 1995, Plant Cell 7:957-970; Shanklin und Cahoon, 1998, Annu. Rev. Plant Physiol. Plant Mol. Biol. 49:611-641; Voelker, 1996, Genetic Engineering, Hrsgb.: JK Setlow, 18:111-13; Gerhardt, 1992, Prog. 45 Lipid R. 31:397-417; Gühnemann-Schäfer & Kindl, 1995, Biochim. Biophys Acta 1256:181-186; Kunau et al., 1995, Prog. Lipid Res. 34:267-342; Stymne et al., 1993, in: Biochemistry and Molecular

## 25

Biology of Membrane and Storage Lipids of Plants, Hrsgb.:  
Murata und Somerville, Rockville, American Society of Plant  
Physiologists, 150-158, Murphy & Ross 1998, Plant Journal.  
13(1):1-16.

5

- Vitamine, Cofaktoren und Nutraceutical wie PUFAs, umfassen eine Gruppe von Molekülen, die höhere Tiere nicht mehr synthetisieren können und somit aufnehmen müssen oder die höheren Tiere nicht mehr ausreichend selbst herstellen können und somit zusätzlich
- 10 aufnehmen müssen, obwohl sie leicht von anderen Organismen, wie Bakterien, synthetisiert werden. Die Biosynthese dieser Moleküle in Organismen, die sie produzieren können, wie in Bakterien, ist im großen und ganzen charakterisiert worden (Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27,
- 15 S. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E., & Packer, L. (1995) "Nutrition, Lipids, Health and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations
- 20 in Malaysia and the Society for Free Radical Research Asia, abgehalten am 1.-3. Sept. 1994 in Penang, Malaysia, AOCS Press, Champaign, IL X, 374 S).

- Die oben erwähnten Moleküle sind entweder selbst biologisch
- 25 aktive Moleküle oder Vorstufen biologisch aktiver Substanzen, die entweder als Elektronenüberträger oder Zwischenprodukte bei einer Vielzahl von Stoffwechselwegen dienen. Diese Verbindungen haben neben ihrem Nährwert auch einen signifikanten industriellen Wert als Farbstoffe, Antioxidantien und Katalysatoren oder andere Ver-
- 30 arbeitungshilfsstoffe. (Einen Überblick über Struktur, Aktivität und industrielle Anwendungen dieser Verbindungen siehe z.B. in Ullmann's Encyclopedia of Industrial Chemistry, "Vitamins", Bd. A27, S. 443-613, VCH: Weinheim, 1996). Mehrfach ungesättigte Fettsäuren haben verschiedene Funktionen und gesundheitsfördernde
- 35 Wirkungen, beispielsweise bei koronarer Herzerkrankung, Entzündungsmechanismen, Kinderernährung usw. Veröffentlichungen und Literaturstellen, einschließlich darin zitierter Literaturstellen, siehe in: Simopoulos, 1999, Am. J. Clin. Nutr. 70 (3. Suppl.):560-569, Takahata et al., Biosci. Biotechnol. Biochem.
- 40 1998, 62(11):2079-2085, Willich und Winther, 1995, Deutsche Medizinische Wochenschrift 120(7):229ff.

## II. Elemente und Verfahren der Erfindung

- 45 Die vorliegende Erfindung beruht unter anderem auf der Entdeckung neuer Moleküle, die hier als Desaturasenukleinsäure- und -proteinmoleküle bezeichnet werden, welche eine Wirkung auf

## 26

die Produktion von Zellmembranen und Lipiden *Phaeodactylum tri-*  
*cornutum* ausüben und beispielsweise die Bewegung von Molekülen  
über diese Membranen beeinflussen. Bei einer Ausführungsform  
nehmen die Desaturasemoleküle am Stoffwechsel von zum Aufbau von  
5 Zellmembranen in Organismen, wie Mikroorganismen und Pflanzen,  
notwendigen Verbindungen teil oder beeinflussen indirekt den  
Transport von Molekülen über diese Membranen. Bei einer bevor-  
zugten Ausführungsform hat die Aktivität der erfindungsgemäßen  
Desaturasemoleküle zur Regulation der Produktion von Membran-  
10 komponenten und des Membrantransports eine Auswirkung auf  
die Produktion der gewünschten Feinchemikalie durch diesen  
Organismus. Bei einer besonders bevorzugten Ausführungsform  
ist die Aktivität der erfindungsgemäßen Desaturasemoleküle  
moduliert, so dass die Ausbeute, Produktion und/oder Effizienz  
15 der Produktion der Stoffwechselwege von Mikroorganismen oder  
Pflanzen, welche die erfindungsgemäßen Desaturasen regulieren,  
moduliert sind und die Effizienz des Transport von Verbindungen  
durch die Membranen verändert ist, was entweder direkt oder  
indirekt die Ausbeute, Produktion und/oder Effizienz der  
20 Produktion einer gewünschten Feinchemikalie durch Mikroorganismen  
und Pflanzen moduliert.

Der Begriff "Desaturase" oder "Desaturasepolypeptid" umfasst  
Proteine, die an der Desaturierung von Fettsäuren teilnehmen.  
25 Beispiele für Desaturasen sind in der SEQ ID NO: 1, 3, 5, 11  
oder ihren Homologen, Derivaten oder Analoga offenbart. Die  
Begriffe Desaturase oder Desaturasenukleinsäuresequenz(en)  
umfassen Nukleinsäuresequenzen, die eine Desaturase kodieren  
und bei denen ein Teil eine kodierende Region und ebenfalls  
30 entsprechende 5'- und 3'-untranslatierte Sequenzbereiche sein  
können. Beispiele für Desaturase-Gene sind die in SEQ ID NO: 1,  
3, 5 oder 11 dargestellten. Die Begriffe Produktion oder  
Produktivität sind im Fachgebiet bekannt und beinhalten die  
Konzentration des Fermentationsproduktes (zum Beispiel der  
35 gewünschten Feinchemikalie), das in einer bestimmten Zeitspanne  
und einem bestimmten Fermentationsvolumen gebildet wird (z.B.  
kg Produkt pro Stunde pro Liter). Der Begriff Effizienz der  
Produktion umfasst die Zeit, die zur Erzielung einer bestimmten  
Produktionsmenge nötig ist (z.B. wie lange die Zelle zur Auf-  
40 richtung einer bestimmten Durchsatzrate einer Feinchemikalie  
benötigt). Der Begriff Ausbeute oder Produkt/Kohlenstoff-  
Ausbeute ist im Fachgebiet bekannt und umfasst die Effizienz  
der Umwandlung der Kohlenstoffquelle in das Produkt (d.h. die  
Feinchemikalie). Dies wird gewöhnlich beispielsweise ausgedrückt  
45 als kg Produkt pro kg Kohlenstoffquelle. Durch Erhöhen der  
Ausbeute oder Produktion der Verbindung wird die Menge der  
gewonnenen Moleküle oder der geeigneten gewonnenen Moleküle

## 27

dieser Verbindung in einer bestimmten Kulturmenge über einen festgelegten Zeitraum erhöht. Die Begriffe Biosynthese oder Biosyntheseweg sind im Fachgebiet bekannt und umfassen die Synthese einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle aus Zwischenverbindungen, beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Die Begriffe Abbau oder Abbauweg sind im Fachgebiet bekannt und umfassen die Spaltung einer Verbindung, vorzugsweise einer organischen Verbindung, durch eine Zelle in Abbauprodukte (allgemeiner gesagt, kleinere oder weniger komplexe Moleküle) beispielsweise in einem Mehrschritt- und stark regulierten Prozess. Der Begriff Stoffwechsel ist im Fachgebiet bekannt und umfasst die Gesamtheit der biochemischen Reaktionen, die in einem Organismus stattfinden. Der Stoffwechsel einer bestimmten Verbindung (z.B. der Stoffwechsel einer Fettsäure) umfasst dann die Gesamtheit der Biosynthese-, Modifikations- und Abbauwege dieser Verbindung in der Zelle, die diese Verbindung betreffen.

Bei einer anderen Ausführungsform können die erfindungsgemäßen Nukleinsäuresequenzen, die für Desaturase-Moleküle codieren, die Produktion eines gewünschten Moleküls, wie einer Feinchemikalie, in einem Mikroorganismus oder in Pflanzen modulieren. Es gibt eine Reihe von Mechanismen, durch die die Veränderung einer erfindungsgemäßen Sequenz die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie aus einem Mikroorganismus- oder Pflanzenstamm, die dieses veränderte Protein enthalten, direkt beeinflussen kann. Die Anzahl oder Aktivität von Desaturasen, die am Transport von Feinchemikalienmolekülen innerhalb oder aus der Zelle beteiligt sind, kann erhöht werden, so dass größere Mengen dieser Verbindungen über Membranen transportiert werden, aus denen sie leichter gewonnen und ineinander umgewandelt werden. Ferner sind Fettsäuren, Triacylglycerine und/oder Lipide selbst wünschenswerte Feinchemikalien; durch Optimierung der Aktivität oder Steigern der Anzahl einer oder mehrerer erfindungsgemäßer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Stören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann es möglich sein, die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmolekülen aus Organismen, wie Mikroorganismen oder Pflanzen, zu erhöhen.

Die Mutagenese der genannten Nukleinsäuresequenzen kann Desaturasen mit veränderten Aktivitäten hervorbringen, welche die Produktion einer oder mehrerer gewünschter Feinchemikalien aus Mikroorganismen oder Pflanzen indirekt beeinflussen. Beispielsweise können Desaturasen, die am Export von Abfallprodukten beteiligt

- sind, eine größere Anzahl oder höhere Aktivität aufweisen, so dass die normalen Stoffwechselabfälle der Zelle (deren Menge möglicherweise aufgrund der Überproduktion der gewünschten Feinchemikalie erhöht ist) effizient exportiert werden, bevor sie die
- 5 Moleküle in der Zelle schädigen können (was die Lebensfähigkeit der Zelle herabsetzen würde) oder die Feinchemikalien-Biosynthesewege stören können (was die Ausbeute, Produktion oder Effizienz der Produktion einer gewünschten Feinchemikalie senken würde). Die relativ großen intrazellulären Mengen der gewünschten Fein-
- 10 chemikalie selbst können ferner für die Zelle toxisch sein, so dass man durch Steigern der Aktivität oder Anzahl von Transportern, die diese Verbindungen aus der Zelle exportieren können, die Lebensfähigkeit der Zelle in Kultur steigern kann, was wiederum zu einer größeren Anzahl an Zellen in der Kultur führt,
- 15 welche die gewünschte Feinchemikalie produzieren. Die Desaturasen können auch so manipuliert werden, dass die entsprechenden Mengen unterschiedlicher Lipid- und Fettsäuremoleküle produziert werden. Dies kann eine erhebliche Auswirkung auf die Lipidzusammensetzung der Zellmembran haben. Da jeder Lipidtyp unterschiedliche physi-
- 20 kalische Eigenschaften aufweist, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität signifikant verändern. Änderungen der Membranfluidität können den Transport von Molekülen über die Membran sowie die Integrität der Zelle beeinflussen, was jeweils eine erhebliche Auswirkung auf die
- 25 Produktion von Feinchemikalien aus Mikroorganismen und Pflanzen in Fermentationskultur im großen Maßstab hat. Pflanzenmembranen verleihen spezifische Eigenschaften, wie Toleranz gegenüber Wärme, Kälte, Salz, Trockenheit sowie Toleranz gegen Pathogene, wie Bakterien und Pilze.
- 30 Die genannten isolierten Nukleinsäuresequenzen sind im Genom eines *Phaeodactylum tricornutum* UTEX646-Stammes enthalten, der über die Algensammlung der University of Texas, Austin verfügbar ist.
- 35 Die Nukleotidsequenz der *Phaeodactylum tricornutum*-cDNA und die abgeleiteten Aminosäuresequenzen der Desaturasen sind in den SEQ ID NO: 1 bis 6 sowie 11 und 12 gezeigt. Es wurden Computeralysen durchgeführt, die diese Nukleotidsequenzen als Sequenzen klassifizieren und/oder identifizieren, die am Stoffwechsel von
- 40 Zellmembrankomponenten beteiligte Proteine oder am Transport von Verbindungen über Zellmembranen beteiligte Proteine bzw. der PUFA Biosynthese codieren. EST's mit der Datenbankeingabe-NO: PT001070010R und PT001078032R durch die Erfinder stellen die erfindungsgemäßen Sequenzen in SEQ ID NO: 1 und 3 dar. Die Sequenz
- 45 des Fragments aus EST PT001070010R wurde ermittelt und ist wie dargestellt in SEQ ID NO: 5. Analog ist die Sequenz des Klonen PT001078032R dargestellt in SEQ ID NO: 1. Den Klonen wurden Gen-

## 29

namen zugewiesen. Abkürzungen bedeuten: Pp = Physcomitrella patens, Pt = Phaeodactylum tricornutum. PT001070010R aus SEQ ID NO: 5 codiert für ein neues Gen homolog zu  $\Delta$ -12-Desaturase und PT001078032R codiert für eine neuartige  $\Delta$ -5-Desaturase. Pt\_des6

5 kann gemäß Beispiel 5a mittels Polymerase Kettenreaktion unter Zuhilfenahme degenerierter Oligonukleotide isoliert werden. Ein so erhaltenes Fragment kann zum Sichten einer cDNA Bank aus Phaeodactylum tricornutum isoliert werden und die codierende Region einer Phaeodactylum tricornutum  $\Delta$ -6-Desaturase erhalten werden.

10 den. Ein so isoliertes Gen wird in Tabelle 1 als Pt\_des6 bezeichnet und ist in SEQ ID NO: 3 dargestellt. Die korrespondierenden Aminosäuresequenzen werden durch Übersetzung des genetischen Codes der Sequenz ID NO: 1, 3 und 5 erhalten und sind als SEQ ID NO: 2, 4 und 6 definiert (siehe auch Tabelle 1). Auch eine weitere Nukleinsäuresequenz, die für eine  $\Delta$ -12-Desaturase codiert, ist Tabelle 1 zu entnehmen. Sie trägt die Klon-Nummer PT001072031R.

Tabelle 1

20

	Gencode	Klonname	Nukleinsäure SEQ ID NO:	Polypeptid SEQ ID NO:
D5 Desaturase	Pt_des5	PT001078032R	1	2
D6 Desaturase	Pt_des6	Pt_des6	3	4
25 D12 Desaturase	Pt_des12	PT001070010R	5	6
D6 Desaturase	Pp_des6	Pp_des6	7	8
D6 Elongase	Pp_PSE1	PP001019019F	9	10
$\Delta$ 12 Desaturase	Pt des12.2	PT001072013R	11	12

30 Die vorliegende Erfindung betrifft auch Proteine mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer Aminosäuresequenz der SEQ ID NO:2, 4, 6 oder 12 ist. Wie hier verwendet, ist ein Protein mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer ausgewählten Aminosäuresequenz ist, zu mindestens etwa 50 % homolog zu der ausgewählten Aminosäuresequenz, z.B. der gesamten ausgewählten Aminosäuresequenz. Ein Protein mit einer Aminosäuresequenz, die im wesentlichen homolog zu einer ausgewählten Aminosäuresequenz ist, kann auch zu mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und stärker

35 bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 % oder 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder mehr homolog zu einer ausgewählten Aminosäuresequenz sein.

45 Die Desaturasen oder der biologisch aktive Teile oder Fragmente davon können am Stoffwechsel von Lipiden zum Aufbau von Zellmembranen oder Speicherlipiden in Organismen teilnehmen und in Kom-

## 30

bination mit weiteren Genen, insbesondere solchen mit Elongaseaktivität zur Elongation von C<sub>18</sub>-bzw C<sub>20-22</sub>-PUFAs benötigten Aktivitäten beitragen, so dass C<sub>18</sub>, C<sub>20</sub>-, C<sub>22</sub>- oder C<sub>24</sub>-PUFAs sowie verwandte PUFAs erhalten werden.

5

Dabei können die Desaturasen in Kombination mit Elongasen und anderen Desaturasen in erfindungsgemäßen Expressionskassetten kloniert werden und zur Transformation von Pflanzen mithilfe von Agrobakterium eingesetzt werden.

10

Verschiedene Aspekte der Erfindung sind eingehender in den folgenden Unterabschnitten beschrieben.

A. Isolierte Nukleinsäuremoleküle

15

Eine Ausführungsform der Erfindung sind isolierte Nukleinsäuren, die von PUFA produzierenden Mikroorganismen stammen und für Polypeptide kodieren, die C<sub>18</sub>-oder C<sub>20-22</sub>-Fettsäuren mit mindestens einer, zwei, drei oder vier Doppelbindungen in der Fettsäure

20 desaturieren.

Eine weitere erfindungsgemäße Ausführungsform sind isolierte Nukleinsäuren, umfassend Nukleotidsequenzen, die für Polypeptide kodieren, die C<sub>18</sub>-bzw C<sub>20</sub>-Fettsäuren mit mindestens ein, zwei, 25 drei oder vier Doppelbindungen in der Fettsäure desaturieren und sind aus der Gruppe, bestehend aus

a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz, 30

b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten 35 Aminosäuresequenzen erhalten werden,

c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, 40 SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf Aminosäureebene aufweisen, ohne dass die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.

45

## 31

Die oben genannte erfindungsgemäße Nukleinsäure stammt von Organismen, wie Ciliaten, Pilzen, Algen oder Dinoflagellaten, die PUFAs synthetisieren können, vorzugsweise von *Phaeodactylum tricornutum* oder nah verwandten Organismen.

- 5 Ein Aspekt der Erfindung betrifft isolierte Expressionskassetten sowie Nukleinsäuremoleküle, die Desaturase-Polypeptide oder biologisch aktive Teile davon kodieren, sowie Nukleinsäurefragmente von diesen, die zur Verwendung als Hybridisierungssonden oder
- 10 Primer zur Identifizierung oder Amplifizierung einer Desaturase-kodierenden Nukleinsäure (z.B. Desaturase-DNA) ausreichen. Der Begriff "Nukleinsäuremolekül", wie hier verwendet, soll DNA-Moleküle (z.B. cDNA oder genomische DNA) und RNA-Moleküle (z.B. mRNA) sowie DNA- oder RNA-Analoga, die mittels Nukleotidanaloga erzeugt
- 15 werden, umfassen. Dieser Begriff umfasst zudem die am 3'- und am 5'-Ende des kodierenden Genbereichs gelegene untranslatierte Sequenz: mindestens 500, bevorzugt 200, besonders bevorzugt 100 Nukleotide der Sequenz stromaufwärts des 5'-Endes des kodierenden Bereichs und mindestens 100, bevorzugt 50, besonders bevorzugt 20
- 20 Nukleotide der Sequenz stromabwärts des 3'-Endes des kodierenden Genbereichs. Das Nukleinsäuremolekül kann einzelsträngig oder doppelsträngig sein, ist aber vorzugsweise doppelsträngige DNA. Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure vorliegen. Eine "isolierte" Nukleinsäure hat vorzugsweise keine Sequenzen, welche die Nukleinsäure in der genomischen DNA des Organismus, aus dem die Nukleinsäure stammt, natürlicherweise flankieren (z.B. Sequenzen, die sich an den 5'- und 3'-Enden der Nukleinsäure befinden). Bei verschiedenen Ausführungsformen
- 25 kann das isolierte Desaturase-Nukleinsäuremolekül zum Beispiel weniger als etwa 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb oder 0,1 kb an Nukleotidsequenzen enthalten, die natürlicherweise das Nukleinsäuremolekül in der genomischen DNA der Zelle, aus der die Nukleinsäure stammt (z.B. eine *Physcomitrella patens*-Zelle)
- 30 flankieren. Ein "isoliertes" Nukleinsäuremolekül, wie ein cDNA-Molekül, kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert
- 35 wird.
- 40

Ein erfindungsgemäße Expressionskassette mit der Struktur SEQ ID NO: 32 -Promotor- SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35 Nukleinsäuremolekül, z.B. ein Nukleinsäuremolekül mit einer Nukleotidsequenz der SEQ ID NO:1 oder eines Teils davon, kann unter

45 Verwendung molekularbiologischer Standardtechniken und der hier bereitgestellten Sequenzinformation isoliert werden. Auch kann

mithilfe von Vergleichsalgorithmen beispielsweise eine homologe Sequenz oder homologe, konservierte Sequenzbereiche auf DNA oder Aminosäureebene identifiziert werden. Beispielsweise kann aus einer *Phaeodactylum tricornutum* cDNA aus einer *Phaeodactylum tricornutum*-Bank isoliert werden, indem die vollständige SEQ ID NO:1, 3, 5 oder 11 oder ein Teil davon als Hybridisierungssonde sowie Standard-Hybridisierungstechniken (wie z.B. beschrieben in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) verwendet werden. Überdies lässt sich ein Nukleinsäuremolekül, umfassend eine vollständige Sequenz der SEQ ID NO: 1, 3, 5 oder 11 oder einen Teil davon, durch Polymerasekettenreaktion isolieren, wobei Oligonukleotidprimer, die auf der Basis dieser Sequenz oder von Teilen davon, insbesondere Regionen um Motive aus Beispiel 5a erstellt werden oder Modifikationen ebensolcher in einzelnen definierten Aminosäuren, verwendet werden (z.B. kann ein Nukleinsäuremolekül, umfassend die vollständige Sequenz der SEQ ID NO:1, 3, 5 oder 11 oder einen Teil davon, durch Polymerasekettenreaktion unter Verwendung von Oligonukleotidprimern isoliert werden, die auf der Basis dieser gleichen Sequenz der SEQ ID NO: 1, 3, 5 oder 11 erstellt worden sind). Zum Beispiel lässt sich mRNA aus Zellen isolieren (z.B. durch das Guanidiniumthiocyanat-Extraktionsverfahren von Chirgwin et al. (1979) *Biochemistry* 18:5294-5299) und cDNA mittels Reverser Transkriptase (z.B. Moloney-MLV-Reverse-Transkriptase, erhältlich von Gibco/BRL, Bethesda, MD, oder AMV-Reverse-Transkriptase, erhältlich von Seikagaku America, Inc., St.Petersburg, FL) herstellen. Synthetische Oligonukleotidprimer zur Amplifizierung mittels Polymerasekettenreaktion lassen sich auf der Basis einer der in SEQ ID NO: 1, 3, 5 oder 11 sowie der in Figur 5a gezeigten Sequenzen oder mithilfe der in SEQ ID NO: 2, 4, 6 oder 12 dargestellten Aminosäuresequenzen erstellen. Eine erfindungsgemäße Nukleinsäure kann unter Verwendung von cDNA oder alternativ von genomischer DNA als Matrize und geeigneten Oligonukleotidprimern gemäß Standard-PCR-Amplifikationstechniken amplifiziert werden. Die so amplifizierte Nukleinsäure kann in einen geeigneten Vektor kloniert werden und mittels DNA-Sequenzanalyse charakterisiert werden. Oligonukleotide, die einer Desaturase-Nukleotidsequenz entsprechen, können durch Standard-Syntheseverfahren, beispielsweise mit einem automatischen DNA-Synthesegerät, hergestellt werden.

Die in SEQ ID NO: 1,3, 5 oder 11 gezeigte cDNA umfasst Sequenzen, die Desaturasen kodieren, (d.h. den "kodierenden Bereich") sowie 5'-untranslatierte Sequenzen und 3'-untranslatierte Sequenzen. Alternativ kann das Nukleinsäuremolekül nur den kodierenden Bereich einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11 um-

## 33

fassen oder kann ganze genomische Fragmente, die aus genomischer DNA isoliert sind, enthalten.

Bei einer weiteren bevorzugten Ausführungsform umfasst ein  
5 erfindungsgemäßes isoliertes Nukleinsäuremolekül ein Nukleinsäuremolekül, das ein Komplement einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen oder eines Teils davon ist. Ein Nukleinsäuremolekül, das zu einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen komplementär ist, ist dann  
10 ausreichend komplementär, wenn es mit einer der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen hybridisieren kann, wodurch ein stabiler Duplex entsteht.

Homologe der neuen Desaturase-Nukleinsäuresequenzen mit der  
15 Sequenz SEQ ID NO: 1, 3, 5 oder 11 bedeutet beispielsweise allelische Varianten mit mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 %, stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 % oder 90 bis 95 % und noch stärker bevorzugt mindestens etwa 95 %, 96 %, 97 %, 98 %, 99 % oder  
20 mehr Homologie zu einer in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen oder ihren Homologen, Derivaten oder Analoga oder Teilen davon. Bei einer weiteren bevorzugten Ausführungsform umfasst ein isoliertes erfindungsgemäßes Nukleinsäuremolekül eine Nukleotidsequenz, die an eine der in SEQ ID NO: 1, 3, 5 oder 11  
25 gezeigten Nukleotidsequenzen oder einen Teil davon hybridisiert, z.B. unter stringenten Bedingungen hybridisiert. Allelische Varianten umfassen insbesondere funktionelle Varianten, die sich durch Deletion, Insertion oder Substitution von Nukleotiden aus/in der in SEQ ID NO: 1, 3, 5 oder 11 dargestellten Sequenz  
30 erhalten lassen, wobei aber die Absicht ist, dass die Enzymaktivität der davon herrührenden synthetisierten Proteine für die Insertion eines oder mehrerer Gene vorteilhafterweise beibehalten wird. Proteine, die noch die enzymatische Aktivität der Desaturase besitzen, das heißt deren Aktivität im wesentlichen  
35 nicht reduziert ist, bedeutet Proteine mit mindestens 10 %, vorzugsweise 20 %, besonders bevorzugt 30 %, ganz besonders bevorzugt 40 % der ursprünglichen Enzymaktivität, verglichen mit dem durch SEQ ID NO: 2, 4, 6 oder 12 kodierten Protein.

40 Homologen der SEQ ID NO: 1, 3, 5 oder 11 bedeuten beispielsweise auch bakterielle, Pilz- und Pflanzenhomologen, verkürzte Sequenzen, einzelsträngige DNA oder RNA der kodierenden und nicht-kodierenden DNA-Sequenz.

45 Homologen der SEQ ID NO: 1, 3, 5 oder 11 bedeutet auch Derivate, wie beispielsweise Promotorvarianten. Die Promotoren stromaufwärts der angegebenen Nukleotidsequenzen können durch

einen oder mehrere Nukleotidaustausche, durch Insertion(en) und/oder Deletion(en) modifiziert werden, ohne dass jedoch die Funktionalität oder Aktivität der Promotoren gestört wird. Es ist weiterhin möglich, dass die Aktivität der Promotoren durch 5 Modifikation ihrer Sequenz erhöht ist oder dass sie vollständig durch aktivere Promotoren, sogar aus heterologen Organismen, ersetzt werden.

Überdies kann das erfindungsgemäße Nukleinsäuremolekül nur 10 einen Teil des kodierenden Bereichs einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11 umfassen, zum Beispiel ein Fragment, das als Sonde oder Primer verwendet werden kann, oder ein Fragment, welches einen biologisch aktiven Abschnitt einer Desaturase kodiert. Die aus der Klonierung des Desaturase-Gens 15 von *Phaeodactylum tricornutum* ermittelten Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von Desaturase-Homologen in anderen Zelltypen und Organismen sowie Desaturase-Homologen aus anderen Mikroalgen oder verwandten Arten gestaltet sind. Die Sonde/der 20 Primer umfasst gewöhnlich im wesentlichen gereinigtes Oligonukleotid. Das Oligonukleotid umfasst gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise etwa 16, stärker bevorzugt etwa 25, 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges einer 25 der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen, eines Antisense-Stranges einer der in SEQ ID NO: 1, 3, 5 oder 11 angegebenen Sequenzen oder seiner Homologen, Derivate oder Analoga oder natürlich vorkommender Mutanten davon hybridisiert. Primer auf der Basis einer Nukleotidsequenz der SEQ ID NO: 1, 3, 5 30 oder 11 können in PCR-Reaktionen zur Klonierung von Desaturase-Homologen verwendet werden. Sonden auf der Basis der Desaturase-Nukleotidsequenzen können zum Nachweis von Transkripten oder genomischen Sequenzen, die das gleiche oder homologe Proteine kodieren, verwendet werden. Bei bevorzugten Ausführungsformen 35 umfasst die Sonde zudem eine daran gebundene Markierungsgruppe, z.B. ein Radioisotop, eine fluoreszierende Verbindung, ein Enzym oder einen Enzym-Cofaktor. Diese Sonden können als Teil eines Test-Kits für genomische Marker zur Identifizierung von Zellen, die eine Desaturase misexprimieren, beispielsweise durch Messen 40 einer Menge einer Desaturase-kodierenden Nukleinsäure in einer Zellenprobe, z.B. Messen der Desaturase-mRNA-Spiegel, oder zur Bestimmung, ob ein genomisches Desaturase-Gen mutiert oder deletiert ist, verwendet werden.

45 Bei einer Ausführungsform kodiert das erfindungsgemäße Nukleinsäuremolekül ein Protein oder einen Teil davon, das/der eine Aminosäuresequenz umfasst, die ausreichend homolog zu einer

## 35

- Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil davon die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über
- 5 diese Membranen teilzunehmen, beibehält. Wie hier verwendet, betrifft der Begriff "ausreichend homolog" Proteine oder Teile davon, deren Aminosäuresequenzen eine minimale Anzahl identischer oder äquivalenter Aminosäurereste (z.B. einen Aminosäurerest mit einer ähnlichen Seitenkette, wie ein Aminosäurerest in
- 10 einer der Sequenzen der SEQ ID NO:2) zu einer Aminosäuresequenz der SEQ ID NO: 2 aufweisen, so dass das Protein oder der Teil davon am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann. Protein-
- 15 bestandteile dieser Stoffwechselwege für Membrankomponenten oder Membrantransportsysteme können, wie hier beschrieben, eine Rolle bei der Produktion und Sekretion einer oder mehrerer Feinchemikalien spielen. Beispiele für diese Aktivitäten sind hier ebenfalls beschrieben. Somit trägt die "Funktion einer
- 20 Desaturase" entweder direkt oder indirekt zur Ausbeute, Produktion und/oder Effizienz der Produktion einer oder mehrerer Feinchemikalien bei. Beispiele für Desaturase-Substratspezifitäten der katalytischen Aktivität sind in Tabelle 5 und 6 angegeben.
- 25 Bei einer weiteren Ausführungsform kodieren Derivate des erfindungsgemäßen Nukleinsäuremoleküls Proteine mit mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder mehr Homologie zu einer vollständigen Aminosäuresequenz der SEQ ID NO:2. Die Homologie der Aminosäuresequenz kann über den gesamten Sequenzbereich mit dem Programm PileUp (J. Mol. Evolution., 25, 351-360, 1987, Higgins et al.,
- 30 CABIOS, 5, 1989:151-153) oder BESTFIT oder GAP bestimmt (Henikoff, S. and Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.)
- 40 Teile von Proteinen, die von den erfindungsgemäßen Desaturase-Nukleinsäuremolekülen kodiert werden, sind vorzugsweise biologisch aktive Teile einer der Desaturasen. Wie hier verwendet, soll der Begriff "biologisch aktiver Teil einer Desaturase", einen Abschnitt, z.B. eine Domäne/ein Motiv, einer Desaturase
- 45 umfassen, der am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann

oder eine in Tabelle 5 und 6 angegebene Aktivität aufweist. Zur Bestimmung, ob eine Desaturase oder ein biologisch aktiver Teil davon am Stoffwechsel von zum Aufbau von Zellmembranen in Mikroorganismen oder Pflanzen notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann, kann ein Test der enzymatischen Aktivität durchgeführt werden. Diese Testverfahren, wie eingehend in Beispiel 8 des Beispielteils beschrieben, sind dem Fachmann geläufig.

- 10 Zusätzliche Nukleinsäurefragmente, die biologisch aktive Abschnitte einer Desaturase kodieren, lassen sich durch Isolierung eines Teils einer der Sequenzen in SEQ ID NO: 1, 3, 5 oder 11, Exprimieren des kodierten Abschnitts der Desaturase oder des Peptids (z.B. durch rekombinante Expression in vitro) und Bestimmen der Aktivität des kodierten Teils der Desaturase oder des Peptids herstellen.

Die Erfindung umfasst zudem Nukleinsäuremoleküle, die sich von einer der in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen (und Teilen davon) aufgrund des degenerierten genetischen Codes unterscheiden und somit die gleiche Desaturase kodieren wie diejenige, die von den in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Nukleotidsequenzen kodiert wird. Bei einer anderen Ausführungsform hat ein erfindungsgemäßes isoliertes Nukleinsäuremolekül eine Nukleotidsequenz, die für ein Protein mit einer in SEQ ID NO: 2, 4, 6 oder 12 gezeigten Aminosäuresequenz kodiert. Bei einer weiteren Ausführungsform kodiert das erfindungsgemäße Nukleinsäuremolekül ein Vollängen-Desaturase-Protein, das zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 (die von einem in SEQ ID NO: 1, 3, 5 oder 11 gezeigten offenen Leseraster kodiert wird) im wesentlichen homolog ist und durch gängige Methoden identifizierbar und isolierbar ist.

Zusätzlich zu den in SEQ ID NO: 1, 3, 5 oder 11 gezeigten Desaturase-Nukleotidsequenzen erkennt der Fachmann, dass DNA-Sequenzpolymorphismen, die zu Änderungen in den Aminosäuresequenzen der Desaturasen führen, innerhalb einer Population (z.B. der *Phaeodactylum tricornutum*-Population) existieren können. Diese genetischen Polymorphismen im Desaturase-Gen können zwischen Individuen innerhalb einer Population aufgrund von natürlicher Variation existieren. Wie hier verwendet, bedeuten die Begriffe "Gen" und "rekombinantes Gen" Nukleinsäuremoleküle mit einem offenen Leseraster, der eine Desaturase, vorzugsweise eine *Phaeodactylum tricornutum*-Desaturase, kodiert. Diese natürlichen Varianten bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz des Desaturase-Gens. Sämtliche und alle dieser Nukleotidvariationen und daraus resultierende

## 37

Aminosäurepolymorphismen in der Desaturase, die das Ergebnis natürlicher Variation sind und die funktionelle Aktivität von Desaturasen nicht verändern, sollen im Umfang der Erfindung enthalten sein.

- 5 Nukleinsäuremoleküle, die den natürlichen Varianten entsprechen, und nicht-*Phaeodactylum tricornutum*-Homologen, -Derivate und -Analoge der *Phaeodactylum tricornutum*-cDNA können auf der Grundlage ihrer Homologie zu der hier offenbarten *Phaeodactylum*
- 10 *tricornutum*-Desaturase-Nukleinsäure unter Verwendung der *Phaeodactylum tricornutum*-cDNA oder eines Teils davon als Hybridisierungssonde gemäß Standard-Hybridisierungstechniken unter stringenten Hybridisierungsbedingungen isoliert werden. Bei einer anderen Ausführungsform ist ein erfindungsgemäßes isoliertes
- 15 Nukleinsäuremolekül mindestens 15 Nukleotide lang und hybridisiert unter stringenten Bedingungen mit dem Nukleinsäuremolekül, das eine Nukleotidsequenz der SEQ ID NO:1, 3, 5 oder 11 umfasst. Bei anderen Ausführungsformen ist die Nukleinsäure mindestens 25, 50, 100, 250 oder mehr Nukleotide lang. Der Begriff "hybridisiert
- 20 unter stringenten Bedingungen", wie hier verwendet, soll Hybridisierungs- und Waschbedingungen beschreiben, unter denen Nukleotidsequenzen, die mindestens 60 % homolog zueinander sind, gewöhnlich aneinander hybridisiert bleiben. Die Bedingungen sind vorzugsweise derart, dass Sequenzen, die mindestens etwa 65 %, 25 stärker bevorzugt mindestens etwa 70 % und noch stärker bevorzugt mindestens etwa 75 % oder stärker zueinander homolog sind, gewöhnlich aneinander hybridisiert bleiben. Diese stringenten Bedingungen sind dem Fachmann bekannt und lassen sich in Current Protocols in Molecular Biology, John Wiley & Sons, N. Y. (1989),
- 30 6.3.1-6.3.6., finden. Ein bevorzugtes, nicht einschränkendes Beispiel für stringente Hybridisierungsbedingungen sind Hybridisierungen in 6 x Natriumchlorid/Natriumcitrat (sodium chloride/sodium citrate = SSC) bei etwa 45°C, gefolgt von einem oder mehreren Waschschritten in 0,2 x SSC, 0,1 % SDS bei 50 bis 65°C.
- 35 Dem Fachmann ist bekannt, dass diese Hybridisierungsbedingungen sich je nach dem Typ der Nukleinsäure und, wenn beispielsweise organische Lösungsmittel vorliegen, hinsichtlich der Temperatur und der Konzentration des Puffers unterscheiden. Die Temperatur unterscheidet sich beispielsweise unter "Standard-Hybridisierungsbedingungen" je nach dem Typ der Nukleinsäure zwischen
- 40 42°C und 58°C in wässrigem Puffer mit einer Konzentration von 0,1 bis 5 x SSC (pH 7,2). Falls organisches Lösungsmittel im oben genannten Puffer vorliegt, zum Beispiel 50 % Formamid, ist die Temperatur unter Standardbedingungen etwa 42°C. Vorzugsweise sind
- 45 die Hybridisierungsbedingungen für DNA:DNA-Hybride zum Beispiel 0,1 x SSC und 20°C bis 45°C, vorzugsweise zwischen 30°C und 45°C. Vorzugsweise sind die Hybridisierungsbedingungen für DNA:RNA-

- Hybride zum Beispiel 0,1 x SSC und 30°C bis 55°C, vorzugsweise zwischen 45°C und 55°C. Die vorstehend genannten Hybridisierungstemperaturen sind beispielsweise für eine Nukleinsäure mit etwa 100 bp (= Basenpaare) Länge und einem G + C-Gehalt von 50 %
- 5 in Abwesenheit von Formamid bestimmt. Der Fachmann weiß, wie die erforderlichen Hybridisierungsbedingungen anhand von Lehrbüchern, wie dem vorstehend erwähnten oder aus den folgenden Lehrbüchern Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989; Hames und Higgins (Hrsgb.) 1985,
- 10 "Nucleic Acids Hybridization: A Practical Approach", IRL Press at Oxford University Press, Oxford; Brown (Hrsgb.) 1991, "Essential Molecular Biology: A Practical Approach", IRL Press at Oxford University Press, Oxford, bestimmt werden können.
- 15 Vorzugsweise entspricht ein erfindungsgemäßes isoliertes Nukleinsäuremolekül, das unter stringenten Bedingungen an eine Sequenz der SEQ ID NO:1, 3, 5 oder 11 hybridisiert, einem natürlich vorkommenden Nukleinsäuremolekül. Wie hier verwendet, betrifft ein "natürlich vorkommendes" Nukleinsäuremolekül ein RNA- oder DNA-
- 20 Molekül mit einer Nukleotidsequenz, die in der Natur vorkommt (z.B. ein natürliches Protein kodiert). Bei einer Ausführungsform kodiert die Nukleinsäure eine natürliche vorkommendes Phyaedactylum tricornutum-Desaturase.
- 25 Zusätzlich zu natürlich vorkommenden Varianten der Desaturase-sequenz, die in der Population existieren können, erkennt der Fachmann ferner, dass auch Änderungen durch Mutation in eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 eingebracht werden können, was zu Änderungen der Aminosäuresequenz der
- 30 kodierten Desaturase führt, ohne dass die Funktionsfähigkeit des Desaturaseproteins beeinträchtigt wird. Beispielsweise lassen sich Nukleotidsusbtitutionen, die an "nicht-essentiellen" Aminosäureresten zu Aminosäuresubstitutionen führen, in einer Sequenz der SEQ ID NO: 2, 4, 6 oder 12 herstellen. Ein "nicht-
- 35 essentieller" Aminosäurerest ist ein Rest, der sich in einer Wildtyp-Desaturasequenz einer der Desaturasen (SEQ ID NO: 2, 4, 6 oder 12) verändern lässt, ohne dass die Aktivität der Desaturase verändert das heißt wesentlich reduziert wird, wohingegen ein "essentieller" Aminosäurerest für die Desaturaseaktivität
- 40 erforderlich ist. Andere Aminosäurereste (z.B. diejenigen, die in der Domäne mit Desaturaseaktivität nicht konserviert oder lediglich semikonserviert sind) können jedoch für die Aktivität nicht essentiell sein und lassen sich somit verändern, ohne dass die Desaturaseaktivität verändert wird.

Folglich betrifft ein weiterer Aspekt der Erfindung Nukleinsäuremoleküle, die Desaturasen kodieren, die veränderte Aminosäurereste enthalten, die für die Desaturaseaktivität nicht essentiell sind. Diese Desaturasen unterscheiden sich in der

5 Aminosäuresequenz von einer Sequenz in SEQ ID NO: 2, 4, 6 oder 12 und behalten dennoch zumindest eine der hier beschriebenen Desaturaseaktivitäten. Das isolierte Nukleinsäuremolekül umfasst bei einer Ausführungsform eine Nukleotidsequenz, die ein Protein kodiert, wobei das Protein eine Aminosäuresequenz mit mindestens

10 etwa 50 % Homologie zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 umfasst und am Stoffwechsel von zum Aufbau von Zellmembranen in *Phaeodactylum tricornutum* notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen kann. Das von dem Nukleinsäuremolekül kodierte Protein ist

15 vorzugsweise mindestens etwa 50 bis 60 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 stärker bevorzugt mindestens etwa 60 bis 70 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 noch stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % homolog zu einer

20 der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12 und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 % oder 99 % homolog zu einer der Sequenzen in SEQ ID NO: 2, 4, 6 oder 12.

Zur Bestimmung der prozentualen Homologie von zwei Aminosäuresequenzen (z.B. einer der Sequenzen der SEQ ID NO: 2, 4, 6 oder

25 sequenzen (z.B. einer der Sequenzen der SEQ ID NO: 2, 4, 6 oder 12 und einer mutierten Form davon) oder von zwei Nukleinsäuren werden die Sequenzen zum Zweck des optimalen Vergleichs untereinander geschrieben (z.B. können Lücken in die Sequenz eines Proteins oder einer Nukleinsäure eingefügt werden, um ein opti-

30 males Alignment mit dem anderen Protein oder der anderen Nukleinsäure zu erzeugen). Die Aminosäurereste oder Nukleotide an den entsprechenden Aminosäurepositionen oder Nukleotidpositionen werden dann verglichen. Wenn eine Position in einer Sequenz (z.B. einer der Sequenzen der SEQ ID NO: 2, 4, 6 oder 12) durch den

35 gleichen Aminosäurerest oder das gleiche Nukleotid wie die entsprechende Stelle in der anderen Sequenz (z.B. einer mutierten Form der aus SEQ ID NO: 2, 4, 6 oder 12 ausgewählten Sequenz) belegt wird, dann sind die Moleküle an dieser Position homolog (d.h. Aminosäure- oder Nukleinsäure-"Homologie", wie hier ver-

40 wendet, entspricht Aminosäure- oder Nukleinsäure-"Identität"). Die prozentuale Homologie zwischen den beiden Sequenzen ist eine Funktion der Anzahl an identischen Positionen, die den Sequenzen gemeinsam sind (d.h. % Homologie = Anzahl der identischen Positionen/Gesamtanzahl der Positionen x 100).

45 Die Begriffe Homologie und Identität sind damit als Synonym anzusehen.

## 40

Ein isoliertes Nukleinsäuremolekül, das eine Desaturase kodiert, die zu einer Proteinsequenz der SEQ ID NO: 2, 4, 6 oder 12 homolog ist, kann durch Einbringen einer oder mehrerer Nukleotidsubstitutionen, -additionen oder -deletionen in eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 erzeugt werden, so dass eine oder mehrere Aminosäuresubstitutionen, -additionen oder -deletionen in das kodierte Protein eingebracht werden. Mutationen können in eine der Sequenzen der SEQ ID NO: 1, 3, 5 oder 11 durch Standardtechniken, wie stellenspezifische Mutagenese und PCR-vermittelte Mutagenese, eingebracht werden. Vorzugsweise werden konservative Aminosäuresubstitutionen an einer oder mehreren der vorhergesagten nicht-essentiellen Aminosäureresten hergestellt. Bei einer "konservativen Aminosäuresubstitution" wird der Aminosäurerest gegen einen Aminosäurerest mit einer ähnlichen Seitenkette ausgetauscht. Im Fachgebiet sind Familien von Aminosäureresten mit ähnlichen Seitenketten definiert worden. Diese Familien umfassen Aminosäuren mit basischen Seitenketten (z.B. Lysin, Arginin, Histidin), sauren Seitenketten (z.B. Asparaginsäure, Glutaminsäure), ungeladenen polaren Seitenketten (z.B. Glycin, Asparagin, Glutamin, Serin, Threonin, Tyrosin, Cystein), unpolaren Seitenketten, (z.B. Alanin, Valin, Leucin, Isoleucin, Prolin, Phenylalanin, Methionin, Tryptophan), beta-verzweigten Seitenketten (z.B. Threonin, Valin, Isoleucin) und aromatischen Seitenketten (z.B. Tyrosin, Phenylalanin, Tryptophan, Histidin). Ein vorhergesagter nicht-essentieller Aminosäurerest in einer Desaturase wird somit vorzugsweise durch einen anderen Aminosäurerest aus der gleichen Seitenkettenfamilie ausgetauscht. Alternativ können bei einer anderen Ausführungsform die Mutationen zufallsgemäß über die gesamte oder einen Teil der Desaturase-kodierenden Sequenz eingebracht werden, z.B. durch Sättigungsmutagenese, und die resultierenden Mutanten können nach der hier beschriebenen Desaturase-Aktivität durchmustert werden, um Mutanten zu identifizieren, die Desaturaseaktivität beibehalten. Nach der Mutagenese einer der Sequenzen der SEQ ID NO: 1, 3, 5 oder 11 kann das kodierte Protein rekombinant exprimiert werden, und die Aktivität des Proteins kann z.B. unter Verwendung der hier beschriebenen Tests (siehe Beispielteil) bestimmt werden.

Zusätzlich zu den Nukleinsäuremolekülen, welche die vorstehend beschriebenen Desaturasen kodieren, betrifft ein weiterer Aspekt der Erfindung isolierte Nukleinsäuremoleküle, die "Antisense" zu den erfindungsgemäßen Nukleinsäuresequenzen sind. Eine "Antisense"-Nukleinsäure umfasst eine Nukleotidsequenz, die zu einer "Sense"-Nukleinsäure, welche ein Protein kodiert, komplementär ist, z.B. komplementär zum kodierenden Strang eines doppelsträngigen cDNA-Moleküls oder komplementär zu einer mRNA-Sequenz.

## 41

- Eine Antisense-Nukleinsäure kann folglich über Wasserstoffbrückenbindungen an eine Sense-Nukleinsäure binden. Die Antisense-Nukleinsäure kann zu einem gesamten Desaturase-kodierenden Strang oder nur zu einem Teil davon komplementär sein. Bei einer
- 5 Ausführungsform ist ein Antisense-Nukleinsäuremolekül "Antisense" zu einem "kodierenden Bereich" des kodierenden Strangs einer Nukleotidsequenz, die eine Desaturase kodiert. Der Begriff "kodierender Bereich" betrifft den Bereich der Nukleotidsequenz, der Codons umfasst, die in Aminosäurereste translatiert werden
- 10 (z.B. den gesamten kodierenden Bereich, der mit dem Stopcodon beginnt und endet, d.h. dem letzten Codon vor dem Stopcodon). Bei einer weiteren Ausführungsform ist das Antisense-Nukleinsäuremolekül "Antisense" zu einem "nicht-kodierenden Bereich" des kodierenden Strangs einer Nukleotidsequenz, die Desaturase
- 15 kodiert. Der Begriff "nicht-kodierender Bereich" betrifft 5'- und 3'-Sequenzen, die den kodierenden Bereich flankieren und nicht in Aminosäuren translatiert werden (d.h. die man auch als 5'- und 3'-untranslatierte Bereiche bezeichnet).
- 20 Unter Voraussetzung der hier offenbarten Desaturase-kodierenden Sequenzen des kodierenden Stranges (z.B. die in SEQ ID NO: 1, 3, 5 oder 11 dargestellten Sequenzen) können erfindungsgemäße Antisense-Nukleinsäuren gemäß den Regeln der Watson-Crick-Basenpaarung gestaltet werden. Das Antisense-Nukleinsäuremolekül kann
- 25 komplementär zum gesamten kodierenden Bereich von Desaturase-mRNA sein, ist aber stärker bevorzugt ein Oligonukleotid, das nur zu einem Teil des kodierenden oder nicht-kodierenden Bereichs von Desaturase-mRNA "Antisense" ist. Das Antisense-Oligonukleotid kann z.B. zu dem Bereich, der die Translationsstartstelle von
- 30 Desaturase-mRNA umgibt, komplementär sein. Ein Antisense-Oligonukleotid kann z.B. etwa 5, 10, 15, 20, 25, 30, 35, 40, 45 oder 50 und mehr Nukleotide lang sein. Eine erfindungsgemäße Antisense-Nukleinsäure kann unter Verwendung chemischer Synthese und enzymatischer Ligationsreaktionen mittels im Fachgebiet bekannter
- 35 Verfahren konstruiert werden. Eine Antisense-Nukleinsäure (z.B. ein Antisense-Oligonukleotid) kann z.B. chemisch synthetisiert werden, wobei natürlich vorkommende Nukleotide oder verschiedentlich modifizierte Nukleotide verwendet werden, die so gestaltet sind, dass sie die biologische Stabilität der Moleküle erhöhen
- 40 oder die physikalische Stabilität des zwischen der Antisense- und der Sense-Nukleinsäure gebildeten Duplexes erhöhen, beispielsweise können Phosphorthioat-Derivate und acridinsubstituierte Nukleotide verwendet werden. Beispiele für modifizierte Nukleotide, die zur Erzeugung der Antisense-Nukleinsäure verwendet
- 45 werden können, sind u.a. 5-Fluoruracil, 5-Bromuracil, 5-Chloruracil, 5-Ioduracil, Hypoxanthin, Xanthin, 4-Acetylcytosin, 5-(Carboxyhydroxymethyl)uracil, 5-Carboxymethylaminomethyl-2-

- thiouridin, 5-Carboxymethylaminomethyluracil, Dihydrouracil, Beta-D-Galactosylqueosin, Inosin, N6-Isopentenyladenin, 1-Methylguanin, 1-Methylinosin, 2,2-Dimethylguanin, 2-Methyladenin, 2-Methylguanin, 3-Methylcytosin, 5-Methylcytosin, N6-Adenin, 5 7-Methylguanin, 5-Methylaminomethyluracil, 5-Methoxyaminomethyl-2-thiouracil, Beta-D-Mannosylqueosin, 5'-Methoxycarboxymethyluracil, 5-Methoxyuracil, 2-Methylthio-N6-isopentyladenin, Uracil-5-oxyessigsäure (v), Wybutoxosin, Desaturaseudouracil, Queosin, 2-Thiocytosin, 5-Methyl-2-thiouracil, 2-Thiouracil, 4-Thiouracil, 10 5-Methyluracil, Uracil-5-oxyessigsäuremethylester, Uracil-5-oxyessigsäure (v), 5-Methyl-2-thiouracil, 3-(3-Amino-3-N-2-carboxypropyl)uracil, (acp3)w und 2,6-Diaminopurin. Die Antisense-Nukleinsäure kann alternativ biologisch unter Verwendung eines Expressionsvektors hergestellt werden, in den eine Nukleinsäure in Antisense-Richtung subkloniert worden ist (d.h. RNA, 15 die von der eingebrachten Nukleinsäure transkribiert wird, ist zu einer Zielnukleinsäure von Interesse in Antisense-Richtung orientiert, was im nachstehenden Unterabschnitt weiter beschrieben ist).
- 20 Die erfindungsgemäßen Antisense-Nukleinsäuremoleküle werden üblicherweise an eine Zelle verabreicht oder in situ erzeugt, so dass sie mit der zellulären mRNA und/oder der genomischen DNA, die eine Desaturase kodiert, hybridisieren oder daran binden, um dadurch die Expression des Proteins, z.B. durch Hemmung der 25 Transkription und/oder Translation, zu hemmen. Die Hybridisierung kann durch herkömmliche Nukleotidkomplementarität unter Bildung eines stabilen Duplexes oder z.B. im Fall eines Antisense-Nukleinsäuremoleküls, das DNA-Duplices bindet, durch spezifische Wechselwirkungen in der großen Furche der Doppelhelix erfolgen.
- 30 Das Antisense-Molekül kann so modifiziert sein, dass es spezifisch an einen Rezeptor oder an ein auf einer ausgewählten Zelloberfläche exprimiertes Antigen bindet, z.B. durch Binden des Antisense-Nukleinsäuremoleküls an ein Peptid oder einen Antikörper, das/der an einen Zelloberflächenrezeptor oder ein Antigen 35 bindet. Das Antisense-Nukleinsäuremolekül kann auch unter Verwendung der hier beschriebenen Vektoren den Zellen zugeführt werden. Zur Erzielung ausreichender intrazellulärer Konzentrationen der Antisense-Moleküle sind Vektorkonstrukte, in denen sich das Antisense-Nukleinsäuremolekül unter der Kontrolle eines 40 starken prokaryotischen, viralen oder eukaryotischen, einschließlich pflanzlichen, Promotors befindet, bevorzugt.

Bei einer weiteren Ausführungsform ist das erfindungsgemäße Antisense-Nukleinsäuremolekül ein  $\alpha$ -anomerer Nukleinsäuremolekül. Ein 45  $\alpha$ -anomerer Nukleinsäuremolekül bildet spezifische doppelsträngige Hybride mit komplementärer RNA, wobei die Stränge im Gegensatz zu gewöhnlichen  $\beta$ -Einheiten parallel zueinander verlaufen. (Gaultier

## 43

et al. (1987) Nucleic Acids Res. 15:6625-6641). Das Antisense-Nukleinsäuremolekül kann zudem ein 2'-o-Methylribonukleotid (Inoue et al. (1987) Nucleic Acids Res. 15:6131-6148) oder ein chimäres RNA-DNA-Analogon (Inoue et al. (1987) FEBS Lett. 5 215:327-330) umfassen.

Bei einer weiteren Ausführungsform ist eine erfindungsgemäße Antisense-Nukleinsäure ein Ribozym. Ribozyme sind katalytische RNA-Moleküle mit Ribonukleaseaktivität, die eine einzelsträngige Nukleinsäure, wie eine mRNA, spalten können, zu der sie einen komplementären Bereich haben. Somit können Ribozyme (z.B. Hammer-head-Ribozyme (beschrieben in Haselhoff und Gerlach (1988) Nature 334:585-591)) zur katalytischen Spaltung von Desaturase-mRNA-Transkripten verwendet werden, um dadurch die Translation von Desaturase-mRNA zu hemmen. Ein Ribozym mit Spezifität für eine Desaturase-kodierende Nukleinsäure kann auf der Basis der Nukleotidsequenz einer der in SEQ ID NO: 1, 3, 5 oder 11 offenbarten Desaturase-cDNA (d.h. oder auf der Basis einer gemäß den in dieser Erfindung gelehrt Verfahren zu isolierenden heterologen Sequenz gestaltet werden. Beispielsweise kann ein Derivat einer Tetrahymena-L-19-IVS-RNA konstruiert werden, wobei die Nukleotidsequenz der aktiven Stelle komplementär zu der Nukleotidsequenz ist, die in einer Desaturase-kodierenden mRNA gespalten werden soll. Siehe z.B. Cech et al., US-Patent Nr. 4,987,071 und Cech et al., US-Patent Nr. 5,116,742. Alternativ kann Desaturase-mRNA zur Selektion einer katalytischen RNA mit einer spezifischen Ribonukleaseaktivität aus einem Pool von RNA-Molekülen verwendet werden. Siehe z.B. Bartel, D., und Szostak, J.W. (1993) Science 261:1411-1418.

Alternativ lässt sich die Desaturase-Gen-Expression hemmen, indem Nukleotidsequenzen, die komplementär zum regulatorischen Bereich einer Desaturase-Nukleotidsequenz (z.B. einem Desaturase-Promotor und/oder -Enhancer) sind, so dirigiert werden, dass Dreifachhelix-Strukturen gebildet werden, welche die Transkription eines Desaturase-Gens in Zielzellen hemmen. Siehe allgemein Helene, C. (1991) Anticancer Drug Res. 6(6) 569-84; Helene, C., et al. (1992) Ann. N. Y. Acad. Sci. 660:27-36; und Maher, L.J. (1992) Bioassays 14(12):807-815.

40 B. Genkonstrukt (= Nukleinsäurekonstrukt, -fragment oder Expressionskassette)

Unter der erfindungsgemäßen Expressionskassette sind die in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 genannten Sequenzen, die sich als Ergebnis des genetischen Codes und/oder deren funktionellen oder nicht funktionellen Derivate

zu verstehen, die mit einem oder mehreren Regulationssignalen vorteilhafterweise zur Erhöhung der Genexpression funktionell verknüpft wurden und welche vorteilhaft die Expression der codierenden Sequenz in der Wirtszelle steuern. Diese regulatorischen

5 Sequenzen sollen die gezielte Expression der Gene und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert und/oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird. Beispielsweise handelt

10 es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen oder anstelle dieser Sequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen

15 noch vorhanden sein und gegebenenfalls genetisch verändert worden sein, so dass die natürliche Regulation ausgeschaltet und die Expression der Gene erhöht wurde. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es wurden keine zusätzlichen Regulationssignale vor die Nukleinsäuresequenz oder

20 dessen Derivate inseriert und der natürliche Promotor mit seiner Regulation wurde nicht entfernt. Stattdessen wurde die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und/oder die Genexpression gesteigert wird. Diese veränderten Promotoren können in Form von Teilsequenzen (= Promotor mit

25 Teilen der erfindungsgemäßen Nukleinsäuresequenzen) auch allein vor das natürliche Gen zur Steigerung der Aktivität gebracht werden. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression

30 der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die  $\Delta$ -5-Desaturase-/ $\Delta$ -6-Desaturase und/oder  $\Delta$ -12-Desaturasegene können in einer oder mehreren Kopien in der Expressionskassette

35 (= Genkonstrukt) enthalten sein.

Die regulatorischen Sequenzen bzw. Faktoren können dabei wie oben beschrieben vorzugsweise die Genexpression der eingeführten Gene positiv beeinflussen und dadurch erhöhen. So kann eine Ver-

40 stärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

## 45

Eine weitere Ausführungsform der Erfindung sind ein oder mehrere Genkonstrukte, die eine oder mehrere Sequenzen enthalten, die durch Seq ID NO: 1, 3, 5, 7, 9 oder 11 definiert sind und gem. SEQ ID NO: 2, 4, 6, 8, 10 oder 12 Polypeptide kodieren. Dabei

5 stammen SEQ ID NO: 1, 3, 5, 7 und 11 von Desaturasen während SEQ ID NO: 9 für eine Elongase codiert. Desaturasen codierende Enzyme, die eine Doppelbindung in  $\Delta$ -5-,  $\Delta$ -6- oder  $\Delta$ -12-Position einführen, wobei das Substrat ein, zwei, drei oder vier Doppel-

10 bindungen aufweisen. Die in SEQ ID NO: 9 dargestellte Sequenz codiert für eine Enzymaktivität, die eine Fettsäure um mindestens zwei Kohlenstoffatome verlängert sowie ihre Homologen, Derivate oder Analoga, die funktionsfähig mit einem oder mehreren Regulationssignalen, vorteilhafterweise zur Steigerung der Genexpression, verbunden sind. Beispiele für diese Regula-

15 tionssequenzen sind Sequenzen, an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäure regulieren. Zusätzlich zu diesen neuen Regulationssequenzen kann die natürliche Regulation dieser Sequenzen vor den eigentlichen Strukturgenen noch vorhanden sein und, wenn geeignet, genetisch modifiziert

20 worden sein, so dass die natürliche Regulation ausgeschaltet worden ist und die Expression der Gene gesteigert worden ist. Das Genkonstrukt kann jedoch auch eine einfachere Struktur haben, d.h. dass keine zusätzlichen Regulationssignale vor der Sequenz SEQ ID NO: 1, 3, 5 oder 11 oder ihren Homologen inseriert worden

25 sind und der natürliche Promotor mit seiner Regulation nicht deletiert worden ist. Statt dessen ist die natürliche Regulationssequenz so mutiert worden, dass keine Regulation mehr stattfindet und die Genexpression verstärkt ist. Das Genkonstrukt kann außerdem vorteilhafterweise eine oder mehrere sogenannte

30 Enhancer-Sequenzen, die funktionsfähig mit dem Promotor verbunden sind und die gesteigerte Expression der Nukleinsäuresequenz ermöglichen, umfassen. Es ist auch möglich, am 3'-Ende der DNA-Sequenzen zusätzlich vorteilhafte Sequenzen zu inserieren, beispielsweise weitere Regulations-elemente oder Terminatoren. Die

35 Desaturasegene und das Elongasegen können im Genkonstrukt in einer oder mehreren Kopien vorliegen. Sie können in einem Genkonstrukt oder mehreren Genkonstrukten vorliegen. Dieses Genkonstrukt oder die Genkonstrukte können zusammen im Wirt-organismus exprimiert werden. Dabei kann das Genkonstrukt oder

40 die Genkonstrukte in einem oder mehreren Vektoren inseriert sein und frei in der Zelle vorliegen oder aber im Genom inseriert sein. Es ist vorteilhaft für die Insertion weiterer Gene in Organismen, wenn weitere Gene im Genkonstrukt vorliegen.

45 Vorteilhafte Regulationssequenzen für das neue Verfahren liegen beispielsweise in Promotoren vor, wie dem cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacI<sup>q</sup>-, T7-, T5-, T3-, gal-, trc-,

- ara-, SP6-,  $\lambda$ -P<sub>R</sub>- oder  $\lambda$ -P<sub>L</sub>-Promotor und werden vorteilhafterweise in Gram-negativen Bakterien angewendet. Weitere vorteilhafte Regulationssequenzen liegen beispielsweise in den Gram-positiven Promotoren amy und SPO2, in den Hefe- oder Pilzpromotoren ADC1, MF $\alpha$ , AC, P-60, CYC1, GAPDH, TEF, rp28, ADH oder in den Pflanzenpromotoren CaMV/35S [Franck et al., Cell 21 (1980) 285-294], PRP1 [Ward et al., Plant. Mol. Biol. 22 (1993)], SSU, OCS, lib4, usp, STLS1, B33, nos oder im Ubiquitin- oder Phaseolin-Promotor vor. In diesem Zusammenhang vorteilhaft sind ebenfalls induzierbare Promotoren, wie die in EP-A-0 388 186 (Benzylsulfonamid-induzierbar), Plant J. 2, 1992:397-404 (Gatz et al., Tetracyclin-induzierbar), EP-A-0 335 528 (Abzisin säure-induzierbar) oder WO 93/21334 (Ethanol- oder Cyclohexenol-induzierbar) beschriebenen Promotoren. Weitere geeignete Pflanzenpromotoren sind der Promotor von cytosolischer FBPase oder der ST-LSI-Promotor der Kartoffel (Stockhaus et al., EMBO J. 8, 1989, 2445), der Phosphoribosylpyrophosphatamidotransferase-Promotor aus Glycine max (Genbank-Zugangsnr. U87999) oder der in EP-A-0 249 676 beschriebene nodienspezifische Promotor. Besonders vorteilhafte Promotoren sind Promotoren, welche die Expression in Geweben ermöglichen, die an der Fettsäurebiosynthese beteiligt sind. Ganz besonders vorteilhaft sind samenspezifische Promotoren, wie der ausführungsgemäße USP Promotor aber auch andere Promotoren wie der LeB4- (Baeumlein et al., Plant J., 1992, 2 (2):233-239), DC3 (Thomas, Plant Cell 1996, 263:359-368), Phaseolin- oder Napin-Promotor. Weitere besonders vorteilhafte Promotoren sind samenspezifische Promotoren, die für monokotyle oder dikotyle Pflanzen verwendet werden können und in US 5,608,152 (Napin-Promotor aus Raps), WO 98/45461 (Oleosin-Promotor aus Arobidopsis), US 5,504,200 (Phaseolin-Promotor aus Phaseolus vulgaris), WO 91/13980 (Bce4-Promotor aus Brassica), von Baeumlein et al., Plant J., 1992, 2 (2):233-239 (LeB4-Promotor aus einer Leguminose) beschrieben sind, wobei sich diese Promotoren für Dikotyledonen eignen. Die folgenden Promotoren eignen sich beispielsweise für Monokotyledonen lpt-2- oder lpt-1-Promotor aus Gerste (WO 95/15389 und WO 95/23230), Hordein-Promotor aus Gerste und andere, in WO 99/16890 beschriebene geeignete Promotoren.
- Es ist im Prinzip möglich, alle natürlichen Promotoren mit ihren Regulationssequenzen, wie die oben genannten, für das neue Verfahren zu verwenden. Es ist ebenfalls möglich und vorteilhaft, zusätzlich synthetische Promotoren zu verwenden.
- Das Genkonstrukt kann, wie oben beschrieben, auch weitere Gene umfassen, die in die Organismen eingebracht werden sollen. Es ist möglich und vorteilhaft, in die Wirtsorganismen Regulationsgene,

- wie Gene für Induktoren, Repressoren oder Enzyme, welche durch ihre Enzymaktivität in die Regulation eines oder mehrerer Gene eines Biosynthesewegs eingreifen, einzubringen und darin zu exprimieren. Diese Gene können heterologen oder homologen Ursprungs sein. Weiterhin können vorteilhaft im Nukleinsäurekonstrukt bzw. Genkonstrukt weitere Biosynthesegene des Fettsäure- oder Lipidstoffwechsels enthalten sein oder aber diese Gene können auf einem weiteren oder mehreren weiteren Nukleinsäurekonstrukten liegen. Vorteilhaft werden als Biosynthesegene des Fettsäure- oder Lipidstoffwechsels ein Gen ausgewählt aus der Gruppe Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym, A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylasen, Lipoxygenasen, Triacylglycerol-Lipasen, Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n) oder deren Kombinationen verwendet.

- Genkonstrukte umfassen vorteilhafterweise zur Expression der anderen vorliegenden Gene weitere 3'- und/oder 5'-terminale Regulationssequenzen zur Steigerung der Expression, die in Abhängigkeit vom gewählten Wirtsorganismus und dem Gen oder den Genen für die optimale Expression ausgewählt werden. Diese Regulationssequenzen sollen, wie oben erwähnt, die spezifische Expression der Gene und die Proteinexpression ermöglichen. Dies kann je nach dem Wirtsorganismus beispielsweise bedeuten, dass das Gen nur nach Induktion exprimiert oder überexprimiert wird oder dass es sofort exprimiert und/oder überexprimiert wird.

- Die Regulationssequenzen oder -faktoren können außerdem vorzugsweise eine vorteilhafte Wirkung auf die Expression der eingebrachten Gene haben und diese somit steigern. Auf diese Weise ist es möglich, dass die Regulationselemente unter Verwendung starker Transkriptionssignale, wie Promotoren und/oder Enhancer, vorteilhafterweise auf Transkriptionsebene verstärkt werden. Es ist jedoch weiterhin auch möglich, die Translation zum Beispiel durch Verbesserung der mRNA-Stabilität zu verstärken.

#### 40 C. Rekombinante Expressionsvektoren und Wirtszellen

- Ein weiterer Aspekt der Erfindung betrifft Vektoren, vorzugsweise Expressionsvektoren, die eine Nukleinsäure enthalten, die eine Desaturase allein (oder einen Teil davon) oder ein unter Punkt b beschriebenes Nukleinsäurekonstrukt in dem die erfindungsgemäße Nukleinsäure allein oder in Kombination mit weiteren Biosynthesegenen des Fettsäure- oder Lipidstoffwechsels

- wie Desaturasen oder Elongasen enthalten ist. Wie hier verwendet, betrifft der Begriff "Vektor" ein Nukleinsäuremolekül, das eine andere Nukleinsäure transportieren kann, an welche es gebunden ist. Ein Vektortyp ist ein "Plasmid", was für eine zirkuläre doppelsträngige DNA-Schleife steht, in die zusätzlichen DNA-Segmente ligiert werden können. Ein weiterer Vektortyp ist ein viraler Vektor, wobei zusätzliche DNA-Segmente in das virale Genom ligiert werden können. Bestimmte Vektoren können in einer Wirtszelle, in die sie eingebracht worden sind, autonom replizieren (z.B. Bakterienvektoren mit bakteriellem Replikationsursprung und episomale Säugervektoren). Andere Vektoren (z.B. nicht-episomale Säugervektoren) werden beim Einbringen in die Wirtszelle in das Genom einer Wirtszelle integriert und dadurch zusammen mit dem Wirtsgenom repliziert.
- 15 Zudem können bestimmte Vektoren die Expression von Genen, mit denen sie funktionsfähig verbunden sind, steuern. Diese Vektoren werden hier als "Expressionsvektoren" bezeichnet. Gewöhnlich haben Expressionsvektoren, die für DNA-Rekombinationstechniken geeignet sind, die Form von Plasmiden. In der vorliegenden Beschreibung können "Plasmid" und "Vektor" austauschbar verwendet werden, da das Plasmid die am häufigsten verwendete Vektorform ist. Die Erfindung soll jedoch diese anderen Expressionsvektorformen, wie virale Vektoren (z.B. replikationsdefiziente Retroviren, Adenoviren und adenoverwandte Viren), die ähnliche Funktionen ausüben, umfassen. Ferner soll der Begriff Vektor auch andere Vektoren, die dem Fachmann bekannt sind, wie Phagen, Viren, wie SV40, CMV, Baculovirus, Adenovirus, Transposons, IS-Elemente, Phasmide, Phagemide, Cosmide, lineare oder zirkuläre DNA, umfassen.
- 30 Die erfindungsgemäßen rekombinanten Expressionsvektoren umfassen eine erfindungsgemäße Nukleinsäure oder ein erfindungsgemäßes Genkonstrukt in einer Form, die sich zur Expression der Nukleinsäure in einer Wirtszelle eignet, was bedeutet, dass die rekombinanten Expressionsvektoren eine oder mehrere Regulationssequenzen, ausgewählt auf der Basis der zur Expression zu verwendenden Wirtszellen, die mit der zu exprimierenden Nukleinsäuresequenz funktionsfähig verbunden ist, umfasst. In einem rekombinanten Expressionsvektor bedeutet "funktionsfähig verbunden", dass die Nukleotidsequenz von Interesse derart an die Regulationssequenz(en) gebunden ist, dass die Expression der Nukleotidsequenz möglich ist und sie aneinander gebunden sind, so dass beide Sequenzen die vorhergesagte, der Sequenz zugeschriebene Funktion erfüllen (z.B. in einem In-vitro-Transkriptions-/Translationssystem oder in einer Wirtszelle, wenn der Vektor in die Wirtszelle eingebracht wird). Der Begriff "Regulationssequenz" soll Promotoren, Enhancer und

andere Expressionskontrollelemente (z.B. Polyadenylierungssignale) umfassen. Diese Regulationssequenzen sind z.B. beschrieben in Goeddel: Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990), oder  
 5 siehe: Gruber und Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, Hrsgb.: Glick und Thompson, Kapitel 7, 89-108, einschließlich der Literaturstellen darin. Regulationssequenzen umfassen solche, welche die konstitutive Expression einer Nukleotidsequenz in vielen Wirtszelltypen steuern, und solche, welche die direkte Expression der  
 10 Nukleotidsequenz nur in bestimmten Wirtszellen unter bestimmten Bedingungen steuern. Der Fachmann weiß, dass die Gestaltung des Expressionsvektors von Faktoren, wie der Auswahl der zu transformierenden Wirtszelle, dem Ausmaß der Expression des  
 15 gewünschten Proteins usw., abhängen kann. Die erfindungsgemäßen Expressionsvektoren können in Wirtszellen eingebracht werden, um dadurch Proteine oder Peptide, einschließlich Fusionsproteinen oder -peptiden, herzustellen, die von den Nukleinsäuren, wie hier beschrieben, kodiert werden (z.B. Desaturasen, mutante Formen von  
 20 Desaturasen, Fusionsproteine usw.).

Die erfindungsgemäßen rekombinanten Expressionsvektoren können zur Expression von Desaturasen und Elongasen in prokaryotischen oder eukaryotischen Zellen gestaltet sein. Beispielsweise können  
 25 Desaturasegene in bakteriellen Zellen, wie *C. glutamicum*, Insektenzellen (unter Verwendung von Baculovirus-Expressionsvektoren), Hefe- und anderen Pilzzellen (siehe Romanos, M.A., et al. (1992) "Foreign gene expression in yeast: a review", Yeast 8:423-488; van den Hondel, C.A.M.J.J., et al. (1991)  
 30 "Heterologous gene expression in filamentous fungi", in: More Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego; und van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics  
 35 of Fungi, Peberdy, J.F., et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge), Algen (Falciatore et al., 1999, Marine Biotechnology.1, 3:239-251), Ciliaten der Typen: Holotrichia, Peritrichia, Spirotrichia, Suctorina, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Desaturase-  
 40 udocohnilembus, Euplotes, Engelmaniella und Stylonychia, insbesondere der Gattung Stylonychia lemnae, mit Vektoren nach einem Transformationsverfahren, wie beschrieben in WO 98/01572, sowie Zellen vielzelliger Pflanzen (siehe Schmidt, R. und Willmitzer, L. (1988) "High efficiency Agrobacterium tumefaciens-mediated  
 45 transformation of Arabidopsis thaliana leaf and cotyledon explants" Plant Cell Rep.:583-586; Plant Molecular Biology and Biotechnology, C Press, Boca Raton, Florida, Kapitel 6/7,

S.71-119 (1993); F.F. White, B. Jené et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press (1993), 128-43; Potrykus, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991), 205-225 (und darin zitierte Literaturstellen)) oder Sugerkzellen exprimiert werden. Geeignete Wirtszellen werden ferner erortert in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Der rekombinante Expressionsvektor kann alternativ, zum Beispiel unter Verwendung von T7-Promotor-Regulationssequenzen und T7-Polymerase, in vitro transkribiert und translatiert werden.

Die Expression von Proteinen in Prokaryoten erfolgt meist mit Vektoren, die konstitutive oder induzierbare Promotoren enthalten, welche die Expression von Fusions- oder nicht-Fusionsproteinen steuern. Fusionsvektoren fugen eine Reihe von Aminosuren an ein darin kodiertes Protein an, gewohnlich am Amino-terminus des rekombinanten Proteins, aber auch am C-Terminus oder fusioniert innerhalb geeigneter Bereiche in den Proteinen. Diese Fusionsvektoren haben gewohnlich drei Aufgaben: 1) die Verstrkung der Expression von rekombinantem Protein; 2) die Erhohung der Loslichkeit des rekombinanten Proteins und 3) die Unterstutzung der Reinigung des rekombinanten Proteins durch Wirkung als Ligand bei der Affinittsreinigung. Bei Fusions-Expressionsvektoren wird oft eine proteolytische Spaltstelle an der Verbindungsstelle der Fusionseinheit und des rekombinanten Proteins eingebracht, so dass die Abtrennung des rekombinanten Proteins von der Fusionseinheit nach der Reinigung des Fusionsproteins mglich ist. Diese Enzyme und ihre entsprechenden Erkennungssequenzen umfassen Faktor Xa, Thrombin und Entero-kinase.

Typische Fusions-Expressionsvektoren sind u.a. pGEX (Pharmacia Biotech Inc; Smith, D.B., und Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) und pRIT5 (Pharmacia, Piscataway, NJ), bei denen Glutathion-S-Transferase (GST), Maltose E-bindendes Protein bzw. Protein A an das rekombinante Zielprotein fusioniert wird. Bei einer Ausfuhrungsform ist die Desaturase-kodierende Sequenz in einen pGEX-Expressionsvektor kloniert, so dass ein Vektor erzeugt wird, der ein Fusionsprotein kodiert, das vom N-Terminus zum C-Terminus GST-Thrombin-Spaltstelle-X-Protein umfasst. Das Fusionsprotein kann durch Affinittschromatographie unter Verwendung von Glutathion-Agarose-Harz gereinigt werden. Rekombinante Desaturase, die nicht an GST fusioniert ist, kann durch Spaltung des Fusionsproteins mit Thrombin gewonnen werden.

## 51

- Beispiele für geeignete induzierbare nicht-Fusions-E. coli-Expressionsvektoren sind u.a. pTrc (Amann et al. (1988) Gene 69:301-315) und pET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 60-89). Die Zielgenexpression vom pTrc-Vektor beruht auf der Transkription durch Wirts-RNA-Polymerase von einem Hybrid-trp-lac-Fusionspromotor. Die Zielgenexpression aus dem pET 11d-Vektor beruht auf der Transkription von einem T7-gnl0-lac-Fusions-Promotor, die von einer coexprimierten viralen RNA-Polymerase (T7 gnl) vermittelt wird. Diese virale Polymerase wird von den Wirtsstämmen BL21 (DE3) oder HMS174 (DE3) von einem residenten  $\lambda$ -Prophagen bereitgestellt, der ein T7 gnl-Gen unter der Transkriptionskontrolle des lacUV 5-Promotors birgt.
- 15 Andere in prokaryotischen Organismen geeignete Vektoren sind dem Fachmann bekannt, diese Vektoren sind beispielsweise in E. coli pLG338, pACYC184, die pBR-Reihe, wie pBR322, die pUC-Reihe, wie pUC18 oder pUC19, die M13mp-Reihe, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III<sup>113</sup>-B1,  $\lambda$ gt11 or pBdCI, 20 in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus PUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667. Eine Strategie zur Maximierung der Expression von rekombinantem Protein ist die Expression des Proteins in einem Wirtsbakterium, dessen Fähigkeit zur proteolytischen Spaltung des rekombinanten 25 Proteins gestört ist (Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Kalifornien (1990) 119-128). Eine weitere Strategie ist die Veränderung der Nukleinsäuresequenz der in einen Expressionsvektor zu inserieren- den Nukleinsäure, so dass die einzelnen Codons für jede Amino- 30 säure diejenigen sind, die vorzugsweise in einem zur Expression ausgewählten Bakterium, wie *C. glutamicum*, verwendet werden (Wada et al. (1992) Nucleic Acids Res. 20:2111-2118). Diese Veränderung der erfindungsgemäßen Nukleinsäuresequenzen erfolgt durch Standard-DNA-Synthesetechniken.
- 35 Bei einer weiteren Ausführungsform ist der Desaturase-Expressionsvektor ein Hefe-Expressionsvektor. Beispiele für Vektoren zur Expression in der Hefe *S. cerevisiae* umfassen pYeDesaturasec1 (Baldari et al. (1987) Embo J. 6:229-234), 40 pMFa (Kurjan und Herskowitz (1982) Cell 30:933-943), pJRY88 (Schultz et al. (1987) Gene 54:113-123) sowie pYES2 (Invitrogen Corporation, San Diego, CA). Vektoren und Verfahren zur Konstruktion von Vektoren, die sich zur Verwendung in anderen Pilzen, wie den filamentösen Pilzen, eignen, umfassen diejenigen, 45 die eingehend beschrieben sind in: van den Hondel, C.A.M.J.J., & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of fungi,

J.F. Peberdy et al., Hrsgb., S. 1-28, Cambridge University Press: Cambridge, oder in: More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, Hrsgb., S. 396-428: Academic Press: San Diego]. Weitere geeignete Hefevektoren sind beispielsweise pAG-1, YEp6, 5 YEp13 oder pEMBLye23.

Alternativ können die erfindungsgemäßen Desaturasen in Insektenzellen unter Verwendung von Baculovirus-Expressionsvektoren exprimiert werden. Baculovirus-Vektoren, die zur Expression von 10 Proteinen in gezüchteten Insektenzellen (z.B. Sf9-Zellen) verfügbar sind, umfassen die pAc-Reihe (Smith et al. (1983) Mol. Cell Biol. 3:2156-2165) und die pVL-Reihe (Lucklow und Summers (1989) Virology 170:31-39).

15 Die oben genannten Vektoren bieten nur einen kleinen Überblick über mögliche geeignete Vektoren. Weitere Plasmide sind dem Fachmann bekannt und sind zum Beispiel beschrieben in: Cloning Vectors (Hrsgb. Pouwels, P.H., et al., Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018).

20

Bei noch einer weiteren Ausführungsform wird eine erfindungsgemäße Nukleinsäure in Säugerzellen unter Verwendung eines Säuger-Expressionsvektors exprimiert. Unter Säugern werden im Sinne der Erfindung alle nicht-humanen Säuger verstanden.

25 Beispiele für Säuger-Expressionsvektoren umfassen pCDM8 (Seed, B. (1987) Nature 329:840) und pMT2PC (Kaufman et al. (1987) EMBO J. 6:187-195). Bei der Verwendung in Säugerzellen werden die Kontrollfunktionen des Expressionsvektors oft von viralen Regulationselementen bereitgestellt. Üblicherweise verwendete 30 Promotoren stammen z.B. aus Polyoma, Adenovirus2, Cytomegalievirus und Simian Virus 40. Weitere geeignete Expressionssysteme für prokaryotische und eukaryotische Zellen siehe in den Kapiteln 16 und 17 von Sambrook, J., Fritsch, E.F., und Maniatis, T., Molecular Cloning: A Laboratory Manual, 2. Auflage, Cold Spring 35 Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Bei einer anderen Ausführungsform kann der rekombinante Säuger-Expressionsvektor die Expression der Nukleinsäure vorzugsweise in 40 einem bestimmten Zelltyp steuern (z.B. werden gewebespezifische Regulationselemente zur Expression der Nukleinsäure verwendet). Gewebespezifische Regulationselemente sind im Fachgebiet bekannt. Nicht beschränkende Beispiele für geeignete gewebespezifische Promotoren sind u.a. der Albuminpromotor (leberspezifisch; 45 Pinkert et al. (1987) Genes Dev. 1:268-277), lymphoidspezifische Promotoren (Calame und Eaton (1988) Adv. Immunol. 43:235-275), insbesondere Promotoren von T-Zellrezeptoren (Winoto und

Baltimore (1989) EMBO J. 8:729-733) und Immunglobulinen (Banerji et al. (1983) Cell 33:729-740; Queen und Baltimore (1983) Cell 33:741-748), neuronspezifische Promotoren (z.B. Neurofilament-Promotor; Byrne und Ruddle (1989) PNAS 86:5473-5477), pankreas-  
5 spezifische Promotoren (Edlund et al., (1985) Science 230:912-916) und milchdrüsenspezifische Promotoren (z.B. Milchserum-Promotor; US-Patent Nr. 4,873,316 und Europäische Patentanmeldung-Veröffentlichung Nr. 264,166). Auch entwicklungsregulierte Promotoren sind umfasst, z.B. die hox-Promotoren  
10 der Maus (Kessel und Gruss (1990) Science 249:374-379) und der Fetoprotein-Promotor (Campes und Tilghman (1989) Genes Dev. 3:537-546).

Bei einer weiteren Ausführungsform können die erfindungsgemäßen  
15 Desaturasen in einzelligen Pflanzenzellen (wie Algen), siehe Falciatore et al., 1999, Marine Biotechnology 1 (3):239-251 und darin zitierte Literaturangaben, und Pflanzenzellen aus höheren Pflanzen (z.B. Spermatophyten, wie Feldfrüchten) exprimiert werden. Beispiele für Pflanzen-Expressionsvektoren  
20 umfassen solche, die eingehend beschrieben sind in: Becker, D., Kemper, E., Schell, J., und Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", Plant Mol. Biol. 20:1195-1197; und Bevan, M.W. (1984) "Binary Agrobacterium vectors for plant transformation",  
25 Nucl. Acids Res. 12:8711-8721; Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.: Kung und R. Wu, Academic Press, 1993, S. 15-38.

30 Eine Pflanzen-Expressionskassette enthält vorzugsweise Regulationssequenzen, welche die Genexpression in Pflanzenzellen steuern können und funktionsfähig verbunden sind, so dass jede Sequenz ihre Funktion, wie Termination der Transkription, erfüllen kann, beispielsweise Polyadenylierungssignale. Bevor-  
35 zugte Polyadenylierungssignale sind diejenigen, die aus Agrobacterium tumefaciens-t-DNA stammen, wie das als Octopinsynthese bekannte Gen 3 des Ti-Plasmids pTiACH5 (Gielen et al., EMBO J. 3 (1984) 835ff.) oder funktionelle Äquivalente davon, aber auch alle anderen in Pflanzen funktionell aktiven Terminatoren  
40 sind geeignet.

Da die Pflanzengenexpression sehr oft nicht auf Transkriptionsebenen beschränkt ist, enthält eine Pflanzen-Expressionskassette vorzugsweise andere funktionsfähig verbunden Sequenzen, wie  
45 Translationenhancer, beispielsweise die Overdrive-Sequenz, welche die 5'-untranslatierte Leader-Sequenz aus Tabakmosaik-

virus, die das Protein/RNA-Verhältnis erhöht, enthält (Gallie et al., 1987, Nucl. Acids Research 15:8693-8711).

Die Pflanzengenexpression muss funktionsfähig mit einem  
5 geeigneten Promotor verbunden sein, der die Genexpression auf  
rechtzeitige, zell- oder gewebespezifische Weise durchführt.  
Bevorzugt sind Promotoren, welche die konstitutive Expression  
herbeiführen (Benfey et al., EMBO J. 8 (1989) 2195-2202), wie  
diejenigen, die von Pflanzenviren stammen, wie 35S CAMV (Franck  
10 et al., Cell 21 (1980) 285-294), 19S CaMV (siehe auch US 5352605  
und WO 84/02913) oder Pflanzenpromotoren, wie der in US 4,962,028  
beschriebene der kleinen Untereinheit der Rubisco.

Andere bevorzugte Sequenzen für die Verwendung zur funktions-  
15 fähigen Verbindung in Pflanzengenexpressions-Kassetten sind  
Targeting-Sequenzen, die zur Steuerung des Genproduktes in sein  
entsprechendes Zellkompartiment notwendig sind (siehe eine Über-  
sicht in Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423  
und darin zitierte Literaturstellen), beispielsweise in die  
20 Vakuole, den Zellkern, alle Arten von Plastiden, wie Amylo-  
plasten, Chloroplasten, Chromoplasten, den extrazellulären Raum,  
die Mitochondrien, das Endoplasmatische Retikulum, Ölkörper,  
Peroxisomen und andere Kompartimente von Pflanzenzellen.

25 Die Pflanzengenexpression lässt sich auch über einen chemisch  
induzierbaren Promotor erleichtern (siehe eine Übersicht in Gatz  
1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48:89-108).  
Chemisch induzierbare Promotoren eignen sich besonders, wenn  
gewünscht wird, dass die Genexpression auf zeitspezifische Weise  
30 erfolgt. Beispiele für solche Promotoren sind ein Salicylsäure-  
induzierbarer Promotor (WO 95/19443), ein Tetracyclin-induzier-  
barer Promotor (Gatz et al. (1992) Plant J. 2, 397-404) und ein  
Ethanol-induzierbarer Promotor.

35 Auch Promotoren, die auf biotische oder abiotische Stress-  
bedingungen reagieren, sind geeignete Promotoren, beispielsweise  
der pathogeninduzierte PRP1-Gen-Promotor (Ward et al., Plant.  
Mol. Biol. 22 (1993) 361-366), der hitzeinduzierbare hsp80-  
Promotor aus Tomate (US 5,187,267), der kälteinduzierbare Alpha-  
40 amylase-Promotor aus Kartoffel (WO 96/12814) oder der durch  
Wunden induzierbare pinII-Promotor (EP-A-0 375 091).

Es sind insbesondere solche Promotoren bevorzugt, welche die  
Genexpression in Geweben und Organen herbeiführen, in denen  
45 die Lipid- und Ölbiosynthese stattfindet, in Samenzellen, wie  
den Zellen des Endosperms und des sich entwickelnden Embryos.  
Geeignete Promotoren sind der Napingen-Promotor aus Raps

- (US 5,608,152), der USP-Promotor aus *Vicia faba* (Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67), der Oleosin-Promotor aus *Arabidopsis* (WO 98/45461), der Phaseolin-Promotor aus *Phaseolus vulgaris* (US 5,504,200), der Bce4-Promotor aus Brassica (WO 91/13980) oder der Legumin-B4-Promotor (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2):233-9) sowie Promotoren, welche die samenspezifische Expression in Monokotyledonen-Pflanzen, wie Mais, Gerste, Weizen, Roggen, Reis usw. herbeiführen. Geeignete beachtenswerte Promotoren sind der lpt2- oder lpt1-Gen-Promotor aus Gerste (WO 95/15389 und WO 95/23230) oder die in WO 99/16890 beschriebenen (Promotoren aus dem Gersten-Hordein-Gen, dem Reis-Glutelin-Gen, dem Reis-Oryzin-Gen, dem Reis-Prolamin-Gen, dem Weizen-Gliadin-Gen, Weizen-Glutelin-Gen, dem Mais-Zein-Gen, dem Hafer-Glutelin-Gen, dem Sorghum-Kasirin-Gen, dem Roggen-Secalin-Gen).

- Insbesondere kann die multiparallele Expression von erfindungsgemäßen Desaturasen allein oder in Kombination mit anderen desaturasen oder Elongasen gewünscht sein. Die Einführung solcher Expressionskassetten kann über eine simultane Transformation mehrerer einzelner Expressionskonstrukte erfolgen oder durch Kombination mehrerer Expressionskassetten auf einem Konstrukt. Auch können mehrere Vektoren mit jeweils mehreren Expressionskassetten transformiert und auf die Wirtszelle übertragen werden.

- Ebenfalls besonders geeignet sind Promotoren, welche die plastidenspezifische Expression herbeiführen, da Plastiden das Kompartiment sind, in dem die Vorläufer sowie einige Endprodukte der Lipidbiosynthese synthetisiert werden. Geeignete Promotoren, wie der virale RNA-Polymerase-Promotor, sind beschrieben in WO 95/16783 und WO 97/06250, und der clpP-Promotor aus *Arabidopsis*, beschrieben in WO 99/46394.

- Die Erfindung stellt zudem einen rekombinanten Expressionsvektor bereit, umfassend ein erfindungsgemäßes DNA Molekül, das in Antisense-Richtung in den Expressionsvektor kloniert ist. d.h. das DNA-Molekül ist derart mit einer regulatorischen Sequenz funktionsfähig verbunden, dass die Expression (durch Transkription des DNA-Moleküls) eines RNA-Moleküls, das zur Desaturase-mRNA "Antisense" ist, ermöglicht wird. Es können Regulationssequenzen ausgewählt werden, die funktionsfähig mit einer in Antisense-Richtung klonierten Nukleinsäure verbunden sind und die kontinuierliche Expression des Antisense-RNA-Moleküls in einer Vielzahl von Zelltypen steuern, zum Beispiel können virale Promotoren und/oder Enhancer oder Regulationssequenzen ausgewählt werden, welche die konstitutive, gewebespezifische

oder zelltypspezifische Expression von Antisense-RNA steuern. Der Antisense-Expressionsvektor kann in Form eines rekombinanten Plasmids, Phagemids oder attenuierten Virus vorliegen, in dem Antisense-Nukleinsäuren unter der Kontrolle eines hochwirksamen regulatorischen Bereichs produziert werden, dessen Aktivität durch den Zelltyp bestimmt werden kann, in den der Vektor eingebracht worden ist. Eine Erläuterung der Regulation der Genexpression mittels Antisense-Genen siehe in Weintraub, H., et al., Antisense-RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Bd. 1(1) 1986.

Ein weiterer Aspekt der Erfindung betrifft Wirtszellen, in die ein erfindungsgemäßer rekombinanter Expressionsvektor eingebracht worden ist. Die Begriffe "Wirtszelle" und "rekombinante Wirtszelle" werden hier untereinander austauschbar verwendet. Selbstverständlich betreffen diese Begriffe nicht nur die bestimmte Zielzelle, sondern auch die Nachkommen oder potentiellen Nachkommen dieser Zelle. Da in aufeinanderfolgenden Generationen aufgrund von Mutation oder Umwelteinflüssen bestimmte Modifikationen auftreten können, sind diese Nachkommen nicht unbedingt mit der Parentalzelle identisch, sind jedoch immer noch vom Umfang des Begriffs, wie hier verwendet, umfasst.

Unter Rekombinant oder Transgen beispielsweise rekombinanten Expressionsvektor oder rekombinanten Wirt oder Wirtszellen im Sinne der Erfindung ist zu verstehen, dass die erfindungsgemäßen Nukleinsäuren und/oder deren natürliche Regulationssequenzen an 5' und 3'-Position der Nukleinsäuren nicht in ihrer natürlichen Umgebung sind, das heißt entweder wurde die Lage der Sequenzen im Herkunftstorganismus verändert oder in diesem wurden die Nukleinsäuresequenzen und/oder die Regulationssequenzen mutiert oder die erfindungsgemäßen Nukleinsäuresequenzen wurden in einen anderen Organismus als den Herkunftstorganismus verbracht oder deren Regulationssequenzen. Auch Kombinationen dieser Veränderungen sind möglich. Unter natürlicher Umgebung ist die Lage einer Nukleinsäuresequenz in einem Organismus zu verstehen, wie er in der Natur vorkommt.

Eine Wirtszelle kann eine prokaryotische oder eukaryotische Zelle sein. Zum Beispiel kann eine Desaturase in Bakterienzellen, wie *C. glutamicum*, Insektenzellen, Pilzzellen oder Säugerzellen (wie Chinesischer Hamster-Ovarzellen (CHO) oder COS-Zellen), Algen, Ciliaten, Pflanzenzellen, Pilzen oder anderen Mikroorganismen, wie *C. glutamicum*, exprimiert werden. Andere geeignete Wirtszellen sind dem Fachmann geläufig.

- Vektor-DNA lässt sich in prokaryotische oder eukaryotische Zellen über herkömmliche Transformations- oder Transfektionstechniken einbringen. Die Begriffe "Transformation" und "Transfektion", Konjugation und Transduktion, wie hier verwendet, sollen eine
- 5 Vielzahl von im Stand der Technik bekannten Verfahren zum Einbringen fremder Nukleinsäure (z.B. DNA) in eine Wirtszelle, einschließlich Calciumphosphat- oder Calciumchlorid-Coprecipitation, DEAE-Dextran-vermittelte Transfektion, Lipofektion, natürliche Kompetenz, chemisch vermittelter Transfer, Elektro-
- 10 poration oder Teilchenbeschuss, umfassen. Geeignete Verfahren zur Transformation oder Transfektion von Wirtszellen, einschließlich Pflanzenzellen, lassen sich finden in Sambrook et al. (Molecular Cloning: A Laboratory Manual., 2. Aufl., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold
- 15 Spring Harbor, NY, 1989) und anderen Labor-Handbüchern, wie Methods in Molecular Biology, 1995, Bd. 44, Agrobacterium protocols, Hrsgb: Gartland und Davey, Humana Press, Totowa, New Jersey.
- 20 Über die stabile Transfektion von Säugerzellen ist bekannt, dass je nach verwendetem Expressionsvektor und verwendeter Transfektionstechnik nur ein kleiner Teil der Zellen die fremde DNA in ihr Genom integriert. Zur Identifikation und Selektion dieser Integrierten wird gewöhnlich ein Gen, das einen selektier-
- 25 baren Marker (z.B. Resistenz gegen Antibiotika) kodiert, zusammen mit dem Gen von Interesse in die Wirtszellen eingebracht. Bevorzugte selektierbare Marker umfassen solche, welche Resistenz gegen Medikamente, wie G418, Hygromycin und Methotrexat, verleihen, oder in Pflanzen solche, welche Resistenz gegen ein
- 30 Herbizid, wie Glyphosphat oder Glufosinat, verleihen. Weitere geeignete Marker sind beispielsweise Marker, welche Gene kodieren, die an Biosynthesewegen von zum Beispiel Zuckern oder Aminosäuren beteiligt sind, wie  $\beta$ -Galactosidase, *ura3* oder *ilv2*. Marker, welche Gene, wie Luziferase, *gfp* oder andere Fluoreszenzgene
- 35 kodieren, sind ebenfalls geeignet. Diese Marker lassen sich in Mutanten verwenden, in denen diese Gene nicht funktionell sind, da sie beispielsweise mittels herkömmlicher Verfahren deletiert worden sind. Ferner können Marker, welche eine Nukleinsäure kodieren, die einen selektierbaren Marker kodiert, in eine Wirts-
- 40 zelle auf dem gleichen Vektor eingebracht werden, wie derjenige, der eine Desaturase kodiert, oder können auf einem gesonderten Vektor eingebracht werden. Zellen, die mit der eingebrachten Nukleinsäure stabil transfiziert worden sind, können zum Beispiel durch Medikamentenselektion identifiziert werden (z.B. überleben
- 45 Zellen, die den selektierbaren Marker integriert haben, wohingegen die anderen Zellen absterben).

## 58

Zur Erzeugung eines homolog rekombinierten Mikroorganismus wird ein Vektor hergestellt, der zumindest einen Abschnitt eines Desaturasegens enthält, in den eine Deletion, Addition oder Substitution eingebracht worden ist, um dadurch das Desaturasegen zu verändern, z.B. funktionell zu disrumpieren. Dieses Desaturasegen ist vorzugsweise ein *Phaeodactylum tricornutum* Desaturasegen, es kann jedoch ein Homologon oder Analogon aus anderen Organismen, sogar aus einer Säuger-, Pilz- oder Insektenquelle verwendet werden. Bei einer bevorzugten Ausführungsform ist der Vektor so gestaltet, dass das endogene Desaturasegen bei homologer Rekombination funktionell disrumpiert wird (d.h. nicht länger ein funktionelles Protein kodiert, auch als Knock-out-Vektor bezeichnet). Alternativ kann der Vektor so gestaltet sein, dass das endogene Desaturasegen bei homologer Rekombination mutiert oder anderweitig verändert wird, aber immer noch ein funktionelles Protein kodiert (z.B. kann der stromaufwärts gelegene regulatorische Bereich so verändert sein, dass dadurch die Expression der endogenen Desaturase verändert wird). Zur Erzeugung einer Punktmutation über homologe Rekombination können auch als Chimeraplasty bekannte DNA-RNA-Hybride verwendet werden, die aus Cole-Strauss et al., 1999, *Nucleic Acids Research* 27(5):1323-1330 und Kmiec, *Gene therapy*, 1999, *American Scientist*, 87(3):240-247 bekannt sind.

Im Vektor für die homologe Rekombination ist der veränderte Abschnitt des Desaturasegens an seinem 5'- und 3'-Ende von zusätzlicher Nukleinsäure des Desaturasegens flankiert, so dass homologe Rekombination zwischen dem exogenen Desaturasegen, das auf dem Vektor vorliegt, und einem endogenen Desaturasegen in einem Mikroorganismus oder einer Pflanze möglich ist. Die zusätzliche flankierende Desaturase-Nukleinsäure ist für eine erfolgreiche homologe Rekombination mit dem endogenen Gen hinreichend lang. Gewöhnlich sind im Vektor mehrere hundert Basenpaare bis zu Kilobasen flankierende DNA (sowohl am 5'- als auch am 3'-Ende) enthalten (eine Beschreibung von Vektoren zur homologen Rekombination siehe z.B. in Thomas, K.R., und Capecchi, M.R. (1987) *Cell* 51:503 oder der Rekombination in *Physcomitrella patens* auf cDNA-Basis in Strepp et al., 1998, *Proc. Natl. Acad. Sci. USA* 95 (8):4368-4373). Der Vektor wird in einen Mikroorganismus oder eine Pflanzenzelle (z.B. mittels Polyethylenglycol-vermittelter DNA) eingebracht, und Zellen, in denen das eingebrachte Desaturasegen mit dem endogenen Desaturasegen homolog rekombiniert ist, werden unter Verwendung im Fachgebiet bekannter Techniken selektiert.

## 59

Bei einer anderen Ausführungsform können rekombinante Organismen, wie Mikroorganismen, hergestellt werden, die ausgewählte Systeme enthalten, welche eine regulierte Expression des eingebrachten Gens ermöglichen. Der Einschluß eines Desaturasegens in einem  
5 Vektor, wobei es unter die Kontrolle des lac-Operons gebracht wird, ermöglicht z.B. die Expression des Desaturasegens nur in Gegenwart von IPTG. Diese Regulationssysteme sind im Fachgebiet bekannt.

- 10 Eine erfindungsgemäße Wirtszelle, wie eine prokaryotische oder eukaryotische Wirtszelle, in Kultur oder auf einem Feld wachsend, kann zur Produktion (d.h. Expression) einer Desaturase verwendet werden. In Pflanzen kann zusätzlich ein alternatives Verfahren durch direkten Transfer von DNA in sich entwickelnde Blüten über
- 15 Elektroporation oder Gentransfer mittels Agrobacterium angewendet werden. Die Erfindung stellt folglich ferner Verfahren zur Produktion von Desaturasen unter Verwendung der erfindungsgemäßen Wirtszellen bereit. Bei einer Ausführungsform umfasst das Verfahren die Anzucht der erfindungsgemäßen Wirtszelle (in die
- 20 ein rekombinanter Expressionsvektor, der eine Desaturase kodiert, eingebracht worden ist, oder in deren Genom ein Gen eingebracht worden ist, das eine Wildtyp- oder veränderte Desaturase kodiert) in einem geeigneten Medium, bis die Desaturase produziert worden ist. Das Verfahren umfasst bei einer weiteren Ausführungsform das
- 25 Isolieren der Desaturasen aus dem Medium oder der Wirtszelle.

- Wirtszellen, die im Prinzip zum Aufnehmen der erfindungsgemäßen Nukleinsäure, des erfindungsgemäßen Genproduktes oder des erfindungsgemäßen Vektors geeignet sind, sind alle
- 30 prokaryotischen oder eukaryotischen Organismen. Die vorteilhafterweise verwendeten Wirtsorganismen sind Organismen, wie Bakterien, Pilze, Hefen, Tier- oder Pflanzenzellen. Weitere vorteilhafte Organismen sind Tiere oder vorzugsweise Pflanzen oder Teile davon. Pilze, Hefen oder Pflanzen werden vorzugsweise ver-
  - 35 wendet, besonders bevorzugt Pilze oder Pflanzen, ganz besonders bevorzugt Pflanzen, wie Ölfruchtpflanzen, die große Mengen an Lipidverbindungen enthalten, wie Raps, Nachtkerze, Canola, Erdnuss, Lein, Soja, Safflor, Sonnenblume, Borretsch, oder Pflanzen, wie Mais, Weizen, Roggen, Hafer, Triticale, Reis,
  - 40 Gerste, Baumwolle, Maniok, Pfeffer, Tagetes, Solanaceen-Pflanzen, wie Kartoffel, Tabak, Aubergine und Tomate, Vicia-Arten, Erbse, Alfalfa, Buschpflanzen (Kaffee, Kakao, Tee), Salix-Arten, Bäume (Ölplume, Kokosnuß) sowie ausdauernde Gräser und Futterfeldfrüchte. Besonders bevorzugte erfindungsgemäße Pflanzen sind
  - 45 Ölfruchtpflanzen, wie Soja, Erdnuss, Raps, Canola, Lein, Nachtkerze, Sonnenblume, Safflor, Bäume (Ölpalme, Kokosnuß).

## D. Isolierte Desaturase

Ein weiterer Aspekt der Erfindung betrifft isolierte Desaturasen und biologisch aktive Teile davon. Ein "isoliertes" oder "ge-  
5 reinigtes" Protein oder ein biologisch aktiver Teil davon ist im wesentlichen frei von zellulärem Material, wenn es durch DNA-Rekombinationstechniken produziert wird, oder von chemischen Vorstufen oder andern Chemikalien, wenn es chemisch synthetisiert wird. Der Begriff "im wesentlichen frei von zellulärem Material"  
10 umfasst Desaturase-Präparationen, in denen das Protein von zellulären Komponenten der Zellen, in denen es natürlich oder rekombinant produziert wird, getrennt ist. Bei einer Ausführungsform umfasst der Ausdruck "im wesentlichen frei von zellulärem Material" Desaturase-Präparationen mit weniger als etwa 30 % (bezogen auf  
15 das Trockengewicht) nicht-Desaturase (hier auch als "verunreinigendes Protein" bezeichnet), stärker bevorzugt weniger als etwa 20 % nicht-Desaturase, noch stärker bevorzugt weniger als etwa 10 % nicht-Desaturase und am stärksten bevorzugt weniger als etwa 5 % nicht-Desaturase. Wenn die Desaturase oder ein biologisch  
20 aktiver Teil davon rekombinant hergestellt worden ist, ist sie/er auch im wesentlichen frei von Kulturmedium, d.h. das Kulturmedium macht weniger als etwa 20 %, stärker bevorzugt weniger als etwa 10 % und am stärksten bevorzugt weniger als etwa 5 % des Volumens der Proteinpräparation aus. Der Begriff "im wesentlichen frei von  
25 chemischen Vorstufen oder anderen Chemikalien" umfasst Desaturase-Präparationen, in denen das Protein von chemischen Vorstufen oder anderen Chemikalien getrennt ist, die an der Synthese des Proteins beteiligt sind. Bei einer Ausführungsform umfasst der Begriff "im wesentlichen frei von chemischen Vorstufen oder  
30 anderen Chemikalien" Desaturase-Präparationen mit weniger als etwa 30 % (bezogen auf das Trockengewicht) chemischen Vorstufen oder nicht-Desaturase-Chemikalien, stärker bevorzugt weniger als etwa 20 % chemischen Vorstufen oder nicht-Desaturase-Chemikalien, noch stärker bevorzugt weniger als etwa 10 % chemischen Vorstufen  
35 oder nicht-Desaturase-Chemikalien und am stärksten bevorzugt weniger als etwa 5 % chemischen Vorstufen oder nicht-Desaturase-Chemikalien. Bei bevorzugten Ausführungsformen weisen isolierte Proteine oder biologisch aktive Teile davon keine verunreinigenden Proteine aus dem gleichen Organismus auf, aus dem die  
40 Desaturase stammt. Diese Proteine werden gewöhnlich durch rekombinante Expression zum Beispiel *Phaeodactylum tricornutum*-Desaturase in Pflanzen wie *Physcomitrella patens* bzw. o.g. oder Mikroorganismen, beispielsweise Bakterien, wie *E. coli*, *Bacillus subtilis*, *C. glutamicum*, Pilzen, wie *Mortierella*, Hefe,  
45 wie *Saccharomyces*, oder Ciliaten wie *Colpidium* oder Algen wie *Phaeodactylum* hergestellt.

## 61

Eine erfindungsgemäße isolierte Desaturase oder ein Teil davon kann auch am Stoffwechsel von zum Aufbau von Zellmembranen in *Phaeodactylum tricornutum* notwendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilnehmen. Bei 5 bevorzugten Ausführungsformen umfasst das Protein oder der Teil davon eine Aminosäuresequenz, die ausreichend homolog zu einer Aminosäuresequenz der SEQ ID NO: 2, 4, 6 oder 12 ist, dass das Protein oder der Teil davon die Fähigkeit, am Stoffwechsel von zum Aufbau von Zellmembranen in *Phaeodactylum tricornutum* not- 10 wendigen Verbindungen oder am Transport von Molekülen über diese Membranen teilzunehmen, beibehält. Der Teil des Proteins ist vorzugsweise ein biologisch aktiver Teil, wie hier beschrieben. Bei einer weiteren bevorzugten Ausführungsform hat eine erfindungsgemäße Desaturase eine der in SEQ ID NO: 2, 4, 6 oder 12 15 gezeigten Aminosäuresequenzen. Bei einer weiteren bevorzugten Ausführungsform hat die Desaturase eine Aminosäuresequenz, die von einer Nukleotidsequenz kodiert wird, die, zum Beispiel unter stringenten Bedingungen, an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert. Bei noch einer weiteren 20 bevorzugten Ausführungsform hat die Desaturase eine Aminosäuresequenz, die von einer Nukleotidsequenz kodiert wird, die mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 %, stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und noch stärker bevorzugt mindestens etwa 96 %, 25 97 %, 98 %, 99 % oder noch homologer zu einer der Aminosäuresequenzen der SEQ ID NO: 2, 4, 6 oder 18 ist. Die erfindungsgemäße bevorzugte Desaturase besitzt vorzugsweise auch mindestens eine der hier beschriebenen Desaturase-Aktivitäten. Zum Beispiel umfasst eine erfindungsgemäße bevorzugte Desaturase eine Amino- 30 säuresequenz, die von einer Nukleotidsequenz kodiert wird, die, zum Beispiel unter stringenten Bedingungen, an eine Nukleotidsequenz der SEQ ID NO: 1, 3, 5 oder 11 hybridisiert und am Stoffwechsel von zum Aufbau von Zellmembranen in *Phaeodactylum tricornutum* notwendigen Verbindungen oder am Transport von Molekülen 35 über diese Membranen teilnehmen kann oder eine Doppelbindung in eine Fettsäure mit ein, zwei, drei oder vier Doppelbindungen und einer Kettenlänge von C<sub>18</sub>, C<sub>20</sub> oder C<sub>22</sub> einführt.

Bei anderen Ausführungsformen ist die Desaturase im wesentlichen 40 homolog zu einer Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 und behält die funktionelle Aktivität des Proteins einer der Sequenzen der SEQ ID NO: 2, 4 oder 6 bei, ihre Aminosäuresequenz unterscheidet sich jedoch aufgrund von natürlicher Variation oder Mutagenese, wie eingehend im obigen Unterabschnitt I beschrieben. 45 Bei einer weiteren Ausführungsform ist die Desaturase folglich ein Protein, das eine Aminosäuresequenz umfasst, die mindestens etwa 50 bis 60 %, vorzugsweise mindestens etwa 60 bis 70 % und

## 62

stärker bevorzugt mindestens etwa 70 bis 80 %, 80 bis 90 %, 90 bis 95 % und am stärksten bevorzugt mindestens etwa 96 %, 97 %, 98 %, 99 % oder noch homologer zu einer vollständigen Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 ist und zumindest eine der

5 hier beschriebenen Desaturase-Aktivitäten aufweist. Bei einer anderen Ausführungsform betrifft die Erfindung ein vollständiges Phaeodactylum tricornutum-Protein, das im wesentlichen homolog zu einer vollständigen Aminosäuresequenz der SEQ ID NO: 2, 4 oder 6 ist.

10

Biologisch aktive Teile einer Desaturase umfassen Peptide, umfassend Aminosäuresequenzen, die von der Aminosäuresequenz einer Desaturase hergeleitet sind, z.B. eine in SEQ ID NO: 2, 4 oder 6 gezeigte Aminosäuresequenz oder die Aminosäuresequenz

15 eines Proteins, das zu einer Desaturase homolog ist, welche weniger Aminosäuren als die Vollängen-Desaturase oder das Vollängenprotein aufweisen, das zu einer Desaturase homolog ist, und zumindest eine Aktivität einer Desaturase aufweisen. Gewöhnlich umfassen biologisch aktive Teile (Peptide, z.B.

20 Peptide, die zum Beispiel 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 oder mehr Aminosäuren lang sind) eine Domäne oder ein Motiv mit mindestens einer Aktivität einer Desaturase. Überdies können andere biologisch aktive Teile, in denen andere Bereiche des Proteins deletiert sind, durch rekombinante

25 Techniken hergestellt und bezüglich einer oder mehrerer der hier beschriebenen Aktivitäten untersucht werden. Die biologisch aktiven Teile einer Desaturase umfassen vorzugsweise ein/eine oder mehrere ausgewählte Domänen/Motive oder Teile davon mit biologischer Aktivität.

30

Desaturasen werden vorzugsweise durch DNA-Rekombinationstechniken hergestellt. Zum Beispiel wird ein das Protein kodierendes Nukleinsäuremolekül in einen Expressionsvektor (wie vorstehend beschrieben) kloniert, der Expressionsvektor wird in eine

35 Wirtszelle (wie vorstehend beschrieben) eingebracht, und die Desaturase wird in der Wirtszelle exprimiert. Die Desaturase kann dann durch ein geeignetes Reinigungsschema mittels Standard-Proteinreinigungstechniken aus den Zellen isoliert werden. Alternativ zur rekombinanten Expression kann eine Desaturase, ein

40 -Polypeptid, oder -Peptid mittels Standard-Peptidsynthesetechniken chemisch synthetisiert werden. Überdies kann native Desaturase aus Zellen (z.B. Endothelzellen) z.B. unter Verwendung eines Anti-Desaturase-Antikörpers isoliert werden, der durch Standardtechniken produziert werden kann, wobei eine erfindungs-

45 gemäßige Desaturase oder ein Fragment davon verwendet wird.

## 63

Die Erfindung stellt auch chimäre Desaturase-Proteine oder Desaturase-Fusionsproteine bereit. Wie hier verwendet, umfasst ein "chimäres Desaturase-Protein" oder "Desaturase-Fusionsprotein" ein Desaturase-Polypeptid, das funktionsfähig an ein nicht-Desaturase-Polypeptid gebunden ist. Ein "Desaturase-Polypeptid" betrifft ein Polypeptid mit einer Aminosäuresequenz, die einer Desaturase entspricht, wohingegen ein "nicht-Desaturase-Polypeptid" ein Polypeptid mit einer Aminosäuresequenz betrifft, die einem Protein entspricht, das im wesentlichen nicht homolog zu der Desaturase ist, z.B. ein Protein, das sich vom der Desaturase unterscheidet und aus dem gleichen oder einem anderen Organismus stammt. Innerhalb des Fusionsproteins soll der Begriff "funktionsfähig verbunden" bedeuten, dass das Desaturase-Polypeptid und das nicht-Desaturase-Polypeptid so miteinander fusioniert sind, dass beide Sequenzen die vorhergesagte, der verwendeten Sequenz zugeschriebene Funktion erfüllen. Das nicht-Desaturase-Polypeptid kann an den N-Terminus oder den C-Terminus des Desaturase-Polypeptids fusioniert sein. Bei einer Ausführungsform ist das Fusionsprotein zum Beispiel ein GST-Desaturase-Fusionsprotein, bei dem die Desaturase-Sequenzen an den C-Terminus der GST-Sequenzen fusioniert sind. Diese Fusionsproteine können die Reinigung der rekombinanten Desaturasen erleichtern. Bei einer weiteren Ausführungsform ist das Fusionsprotein eine Desaturase, die eine heterologe Signalsequenz an ihrem N-Terminus aufweist. In bestimmten Wirtszellen (z.B. Säuger-Wirtszellen) kann die Expression und/oder Sekretion einer Desaturase durch Verwendung einer heterologen Signalsequenz gesteigert werden.

Ein erfindungsgemäßes chimäres Desaturase-Protein oder Desaturase-Fusionsprotein wird durch Standard-DNA-Rekombinationstechniken hergestellt. Zum Beispiel werden DNA-Fragmente, die unterschiedliche Polypeptidsequenzen kodieren, gemäß herkömmlicher Techniken im Leseraster aneinander ligiert, indem beispielsweise glatte oder überhängende Enden zur Ligation, Restriktionsenzymsspaltung zur Bereitstellung geeigneter Enden, Auffüllen kohäsiver Enden, wie erforderlich, Behandlung mit alkalischer Phosphatase, um ungewollte Verknüpfungen zu vermeiden, und enzymatische Ligation eingesetzt werden. Bei einer weiteren Ausführungsform kann das Fusionsgen durch herkömmliche Techniken, einschließlich DNA-Syntheseautomaten, synthetisiert werden. Alternativ kann eine PCR-Amplifizierung von Genfragmenten unter Verwendung von Ankerprimern durchgeführt werden, die komplementäre Überhänge zwischen aufeinanderfolgenden Genfragmenten erzeugen, die anschließend miteinander hybridisiert und reamplifiziert werden können, so dass eine chimäre Gensequenz erzeugt wird (siehe zum Beispiel Current Protocols in Molecular

Biology, Hrsgb. Ausubel et al., John Wiley & Sons: 1992). Überdies sind viele Expressionsvektoren kommerziell erhältlich, die bereits eine Fusionseinheit (z.B. ein GST-Polypeptid) kodieren. Eine Desaturase-kodierende Nukleinsäure kann in einen solchen  
5 Expressionsvektor kloniert werden, so dass die Fusionseinheit im Leseraster mit dem Desaturase-Protein verbunden ist.

Homologe der Desaturase können durch Mutagenese, z.B. durch spezifische Punktmutation oder Verkürzung der Desaturase, erzeugt  
10 werden. Der Begriff "Homologe", wie hier verwendet, betrifft eine variante Form der Desaturase, die als Agonist oder Antagonist der Desaturase-Aktivität wirkt. Ein Agonist der Desaturase kann im wesentlichen die gleiche Aktivität wie die oder einen Teil der biologischen Aktivitäten der Desaturase beibehalten. Ein  
15 Antagonist der Desaturase kann eine oder mehrere Aktivitäten der natürlich vorkommenden Form der Desaturase durch zum Beispiel kompetitive Bindung an ein stromabwärts oder -aufwärts gelegenes Element der Stoffwechselkaskade für Zellmembrankomponenten, welche die Desaturase umfasst, oder durch Bindung an eine  
20 Desaturase, welche den Transport von Verbindungen über Zellmembranen vermittelt, hemmen, wodurch die Translokation gehemmt wird.

Bei einer alternativen Ausführungsform können Homologe der Desaturase durch Sichten kombinatorischer Banken von Mutanten,  
25 z.B. Verkürzungsmutanten, der Desaturase hinsichtlich Desaturase-Agonisten- oder -Antagonisten-Aktivität identifiziert werden. Bei einer Ausführungsform wird eine variierte Bank von Desaturase-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt und durch eine variierte Genbank kodiert. Eine  
30 variierte Bank von Desaturase-Varianten kann z.B. durch enzymatische Ligation eines Gemisches von synthetischen Oligonukleotiden in Gensequenzen hergestellt werden, so dass sich ein degenerierter Satz potentieller Desaturase-Sequenzen als individuelle Polypeptide oder alternativ als Satz größerer  
35 Fusionsproteine (z.B. für das Phage-Display), die diesen Satz von Desaturase-Sequenzen enthalten, exprimieren lässt. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Desaturase-Homologen aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische  
40 Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt und das synthetische Gen dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Satzes von Genen ermöglicht die Bereitstellung sämtlicher Sequenzen, die den gewünschten Satz  
45 an potentiellen Desaturase-Sequenzen kodieren, in einem Gemisch. Verfahren zur Synthese degenerierter Oligonukleotide sind im Fachgebiet bekannt (siehe z.B. Narang, S.A. (1983) Tetrahedron

39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

- 5 Zusätzlich können Banken von Desaturase-Fragmenten zur Herstellung einer variierten Population von Desaturase-Fragmenten für das Sichten und für die anschließende Selektion von Homologen einer Desaturase verwendet werden. Bei einer Ausführungsform kann eine Bank von Fragmenten der kodierenden Sequenz durch
- 10 Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Desaturase-Sequenz mit einer Nuklease unter Bedingungen, unter denen Doppelstrangbrüche nur etwa einmal pro Molekül erfolgen, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, welche Sense/Antisense-Paare von
- 15 verschiedenen Produkten mit Doppelstrangbrüchen umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuklease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Mit diesem Verfahren kann eine Expressionsbank hergeleitet werden,
- 20 die N-terminale, C-terminale und interne Fragmente der Desaturase verschiedener Größen kodiert.

- Im Fachgebiet sind mehrere Techniken für das Sichten von Genprodukten in kombinatorischen Banken, die durch Punktmutationen
- 25 oder Verkürzung hergestellt worden sind, und für das Sichten von cDNA-Banken nach Genprodukten mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Sichtung der Genbanken anpassen, die durch kombinatorische Mutagenese von Desaturase-Homologen erzeugt worden sind. Die am häufigsten
- 30 verwendeten Techniken zum Sichtung großer Genbanken, die einer Analyse mit hohem Durchsatz unterworfen werden können, umfassen gewöhnlich das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren von geeigneten Zellen mit der resultierenden Vektorenbank und Expressieren der kombinatorischen
- 35 Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen kodiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine neue Technik, die die Häufigkeit funktioneller Mutanten in den Banken erhöht, kann in Kombination mit den
- 40 Sichtungstests zur Identifikation von Desaturase-Homologen verwendet werden (Arkin und Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).
- 45 Eine weitere bekannte Technik zur Veränderung von katalytischen Eigenschaften von Enzymen bzw. deren codierenden Genen ist das "Gen-Shuffling" (siehe z.B. in Stemmer, PNAS 1994, 91:

10747-10751, WO9720078 oder WO9813487), das eine Kombination von Genfragmenten darstellt, wobei diese Neukombination zusätzlich noch durch fehlerhafte Polymerasekettenreaktionen variiert werden kann und somit eine hohe zu testende Sequenzdiversität schafft.

- 5 Voraussetzung für den Einsatz eines solchen Ansatzes ist jedoch ein geeignetes Screeningsystem, um die erstellte Gendiversität auf Funktionalität zu überprüfen.

Insbesondere für die Sichtung von Desaturaseaktivitäten

- 10 ist ein Sichtungsverfahren Voraussetzung, das PUFA-abhängig Enzymaktivität(en) erfaßt. Bzgl. Desaturaseaktivitäten mit Spezifität für PUFAs kann man in *Mucor*-Species, die durch bekannte Transformationsverfahren mit gewünschten Genkonstrukten transformierbar sind, die Toxizität von Arachidonsäure in An-
- 15 wesenheit eines toxischen Metaboliten (hier: Salicylsäure oder Salicylsäurederivate) nutzen (Eroshin et al., *Mikrobiologiya*, Vol. 65, No.1 1996, Seiten 31-36), um eine wachstumsbasierte Erstsichtung durchzuführen. Resultierende Klone können dann einer Analyse ihrer Lipidinhaltstoffe mittels Gaschromatographie und
- 20 Massenspektroskopie unterzogen werden, um Edukte und Produkte in Art und Menge zu erfassen.

Bei einer weiteren Ausführungsform können Tests auf Zellbasis zur Analyse einer variierten Desaturase-Bank unter Verwendung

- 25 von weiteren im Fachgebiet bekannten Verfahren ausgenutzt werden.

#### E. Erfindungsgemäße Verwendungen und Verfahren

Die hier beschriebenen Nukleinsäuremoleküle, Proteine, Protein-

30 homologen, Fusionsproteine, Primer, Vektoren und Wirtszellen können bei einem oder mehreren der nachstehenden Verfahren verwendet werden: Identifikation von *Phaeodactylum* und verwandten Organismen, Kartierung der Genome von Organismen, die mit *Phaeodactylum tricornutum* verwandt sind, Identifikation und

- 35 Lokalisierung von *Phaeodactylum tricornutum*-Sequenzen von Interesse, Evolutionsstudien, Bestimmung von Desaturase-Proteinbereichen, die für die Funktion notwendig sind, Modulation einer Desaturase-Aktivität; Modulation des Stoffwechsels einer oder mehrerer Zellmembrankomponenten; Modulation des Transmembran-
- 40 transports einer oder mehrerer Verbindungen sowie Modulation der zellulären Produktion einer gewünschten Verbindung, wie einer Feinchemikalie. Die erfindungsgemäßen Desaturase-Nukleinsäuremoleküle haben eine Vielzahl von Verwendungen. Sie können zunächst zur Identifikation eines Organismus als *Phaeodactylum*
- 45 *tricornutum* oder als naher Verwandter davon verwendet werden. Sie können auch zur Identifikation des Vorliegens von *Phaeodactylum tricornutum* oder eines Verwandten davon in einer Mischpopulation

- von Mikroorganismen verwendet werden. Die Erfindung stellt die Nukleinsäuresequenzen einer Reihe von *Phaeodactylum tricornutum*-Genen bereit; durch Sondieren der extrahierten genomischen DNA einer Kultur einer einheitlichen oder gemischten Population von
- 5 Mikroorganismen unter stringenten Bedingungen mit einer Sonde, die einen Bereich eines *Phaeodactylum tricornutum* -Gens oder von Teilen davon überspannt, das für diesen Organismus einzigartig ist, kann man bestimmen, ob dieser Organismus vorliegt. *Phaeodactylum tricornutum* selbst werden zur kommerziellen Produktion
- 10 mehrfach ungesättigter Säuren verwendet und eignen darüber hinaus zur PUFA-Produktion auch in anderen Organismen insbesondere wenn erreicht werden soll, dass resultierende PUFAs auch in die Triacylglycerolfraktion eingebaut werden sollen.
- 15 Ferner können die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle als Marker für spezifische Bereiche des Genoms dienen. Dies ist nicht nur zur Kartierung des Genoms, sondern auch für funktionelle *Phaeodactylum tricornutum*-Proteinen geeignet. Zur Identifikation des Genombereichs, an den ein bestimmtes DNA-
- 20 bindendes Protein von *Phaeodactylum tricornutum* bindet, könnte das *Phaeodactylum tricornutum*-Genom zum Beispiel gespalten werden und die Fragmente mit dem DNA-bindenden Protein inkubiert werden. Diejenigen, die das Protein binden, können zusätzlich mit den erfindungsgemäßen Nukleinsäuremolekülen, vorzugsweise mit leicht
- 25 nachweisbaren Markierungen, sondiert werden; die Bindung eines solchen Nukleinsäuremoleküls an das Genomfragment ermöglicht die Lokalisierung des Fragments auf der Genomkarte von *Phaeodactylum tricornutum* und erleichtert, wenn dies mehrmals mit unterschiedlichen Enzymen durchgeführt wird, eine rasche Bestimmung der
- 30 Nukleinsäuresequenz, an die das Protein bindet. Die erfindungsgemäßen Nukleinsäuremoleküle können zudem ausreichend homolog zu den Sequenzen verwandter Arten sein, dass diese Nukleinsäuremoleküle als Marker für die Konstruktion einer genomischen Karte bei verwandten Pilzen oder Algen dienen können.
- 35 Die erfindungsgemäßen Desaturase-Nukleinsäuremoleküle eignen sich auch für Evolutions- und Proteinstruktur-Untersuchungen. Die Stoffwechsel- und Transportprozesse, an denen die erfindungsgemäßen Moleküle beteiligt sind, werden von vielen
- 40 prokaryotischen und eukaryotischen Zellen genutzt; durch Vergleich der Sequenzen der erfindungsgemäßen Nukleinsäuremoleküle mit solchen, die ähnliche Enzyme aus anderen Organismen kodieren, kann der Evolutions-Verwandschaftsgrad der Organismen bestimmt werden. Entsprechend ermöglicht ein solcher Vergleich
- 45 die Bestimmung, welche Sequenzbereiche konserviert sind und welche nicht, was bei der Bestimmung von Bereichen des Proteins hilfreich sein kann, die für die Enzymfunktion essentiell sind.

Dieser Typ der Bestimmung ist für Proteinengineering-Untersuchungen wertvoll und kann einen Hinweis darauf geben, wieviel Mutagenese das Protein tolerieren kann, ohne die Funktion zu verlieren.

5

Die Manipulation der erfindungsgemäßen Desaturase-Nukleinsäuremoleküle kann zur Produktion von Desaturasen mit funktionellen Unterschieden zu den Wildtyp-Desaturasen führen. Die Effizienz oder Aktivität dieser Proteine kann verbessert sein, sie können  
10 in größeren Anzahlen als gewöhnlich in der Zelle zugegen sein, oder ihre Effizienz oder Aktivität kann verringert sein. Verbesserte Effizienz oder Aktivität bedeutet beispielsweise, dass das Enzym eine höhere Selektivität und/oder Aktivität, vorzugsweise eine mindestens 10 % höhere, besonders bevorzugt eine  
15 mindestens 20 % höhere Aktivität, ganz besonders bevorzugt eine mindestens 30 % höhere Aktivität als das ursprüngliche Enzym aufweist.

Es gibt eine Reihe von Mechanismen, durch die die Veränderung  
20 einer erfindungsgemäßen Desaturase die Ausbeute, Produktion und/oder Effizienz der Produktion einer Feinchemikalie, welche ein solches verändertes Protein enthält, direkt beeinflussen kann. Die Gewinnung von Feinchemikalien-Verbindungen aus Kulturen von Ciliaten, Algen oder Pilzen im großen Maßstab ist signifi-  
25 kant verbessert, wenn die Zelle die gewünschten Verbindungen sezerniert, da diese Verbindungen aus dem Kulturmedium (im Gegensatz zur Extraktion aus der Masse der gezüchteten Zellen) leicht gereinigt werden können. Ansonsten lässt sich die Reinigung verbessern, wenn die Zelle in vivo Verbindungen in einem speziali-  
30 sierten Kompartiment mit einer Art Konzentrationsmechanismus speichert. Bei Pflanzen, die Desaturasen exprimieren, kann ein gesteigerter Transport zu besserer Verteilung innerhalb des Pflanzengewebes und der -organe führen. Durch Vergrößern der Anzahl oder der Aktivität von Transportermolekülen, welche Fein-  
35 chemikalien aus der Zelle exportieren, kann es möglich sein, die Menge der produzierten Feinchemikalie, die im extrazellulären Medium zugegen ist, zu steigern, wodurch Ernte und Reinigung oder bei Pflanzen eine effizientere Verteilung erleichtert werden. Zur effizienten Überproduktion einer oder mehrerer Feinchemikalien  
40 sind dagegen erhöhte Mengen an Cofaktoren, Vorläufermolekülen und Zwischenverbindungen für die geeigneten Biosynthesewege erforderlich. Durch Vergrößern der Anzahl und/oder der Aktivität von Transporterproteinen, die am Import von Nährstoffen, wie Kohlenstoffquellen (d.h. Zuckern), Stickstoffquellen (d.h. Aminosäuren,  
45 Ammoniumsalzen), Phosphat und Schwefel, beteiligt sind, kann man die Produktion einer Feinchemikalie aufgrund der Beseitigung aller Einschränkungen des Nährstoffangebots beim Biosynthese-

prozess verbessern. Fettsäuren, wie PUFAs, und Lipide, die PUFAs enthalten, sind selbst wünschenswerte Feinchemikalien; durch Optimieren der Aktivität oder Erhöhen der Anzahl einer oder mehrerer erfindungsgemäßer Desaturasen, die an der Biosynthese dieser Verbindungen beteiligt sind, oder durch Stören der Aktivität einer oder mehrerer Desaturasen, die am Abbau dieser Verbindungen beteiligt sind, kann man somit die Ausbeute, Produktion und/oder Effizienz der Produktion von Fettsäure- und Lipidmoleküle in Ciliaten, Algen, Pflanzen, Pilzen, Hefen oder anderen Mikroorganismen steigern.

Die Manipulation eines oder mehrerer erfindungsgemäßer Desaturase-Gene kann ebenfalls zu Desaturasen mit veränderten Aktivitäten führen, welche die Produktion einer oder mehrerer gewünschter Feinchemikalien aus Algen, Pflanzen, Ciliaten oder Pilzen indirekt beeinflussen. Die normalen biochemischen Stoffwechselprozesse führen z.B. zur Produktion einer Vielzahl an Abfallprodukten (z.B. Wasserstoffperoxid und andere reaktive Sauerstoffspezies), die diese Stoffwechselprozesse aktiv stören können (z.B. nitriert Peroxynitrit bekanntlich Tyrosin-Seitenketten, wodurch einige Enzyme mit Tyrosin im aktiven Zentrum inaktiviert werden (Groves, J.T. (1999) Curr. Opin. Chem. Biol. 3(2);226-235)). Diese Abfallprodukte werden zwar üblicherweise ausgeschieden, aber die zur fermentativen Produktion im großen Maßstab verwendeten Zellen werden für die Überproduktion einer oder mehrerer Feinchemikalien optimiert und können somit mehr Abfallprodukte produzieren als für eine Wildtypzelle üblich. Durch Optimieren der Aktivität einer oder mehrerer erfindungsgemäßer Desaturasen, die am Export von Abfallmolekülen beteiligt sind, kann man die Lebensfähigkeit der Zelle verbessern und eine effiziente Stoffwechselaktivität aufrechterhalten. Auch das Vorliegen hoher intrazellulärer Mengen der gewünschten Feinchemikalie kann tatsächlich für die Zelle toxisch sein, so dass man durch Steigern der Fähigkeit der Zelle zur Sekretion dieser Verbindungen die Lebensfähigkeit der Zelle verbessern kann.

Die erfindungsgemäßen Desaturasen können ferner so manipuliert sein, dass die relativen Mengen verschiedener Lipid- und Fettsäuremoleküle verändert werden. Dies kann eine entscheidende Auswirkung auf die Lipidzusammensetzung der Zellmembran haben. Da jeder Lipidtyp unterschiedliche physikalische Eigenschaften hat, kann eine Veränderung der Lipidzusammensetzung einer Membran die Membranfluidität signifikant verändern. Änderungen der Membranfluidität können den Transport von Molekülen über die Membran beeinflussen, was, wie vorstehend erläutert, den Export von Abfallprodukten oder der produzierten Feinchemikalie oder den Import notwendiger Nährstoffe modifizieren kann. Diese Änderungen

der Membranfluidität können auch die Integrität der Zelle entscheidend beeinflussen; Zellen mit vergleichsweise schwächeren Membranen sind anfälliger gegenüber abiotischen und biotischen Stressbedingungen, welche die Zelle beschädigen oder abtöten  
5 können. Durch Manipulieren von Desaturasen, die an der Produktion von Fettsäuren und Lipiden für den Membranaufbau beteiligt sind, so dass die resultierende Membran eine Membranzusammensetzung hat, die für die in den Kulturen, die zur Produktion von Feinchemikalien verwendet werden, herrschenden Umweltbedingungen  
10 empfänglicher sind, sollte ein größerer Anteil der Zellen überleben und sich vermehren. Größere Mengen an produzierenden Zellen sollten sich in größeren Ausbeuten, höherer Produktion oder Effizienz der Produktion der Feinchemikalie aus der Kultur manifestieren.

15

Die vorstehend genannten Mutagenesestrategien für Desaturasen, die zu erhöhten Ausbeuten einer Feinchemikalie führen sollen, sollen nicht beschränkend sein; Variationen dieser Strategien sind dem Fachmann leicht ersichtlich. Unter Verwendung dieser  
20 Mechanismen und mithilfe der hier offenbarten Mechanismen können die erfindungsgemäßen Nukleinsäure- und Proteinmoleküle zur Erzeugung von Algen, Ciliaten, Pflanzen, Tieren, Pilzen oder anderen Mikroorganismen, wie *C. glutamicum*, verwendet werden, die mutierte Desaturase-Nukleinsäure- und Proteinmoleküle expri-  
25 mieren, so dass die Ausbeute, Produktion und/oder Effizienz der Produktion einer gewünschten Verbindung verbessert wird. Diese gewünschte Verbindung kann ein beliebiges natürliches Produkt von Algen, Ciliaten, Pflanzen, Tieren, Pilzen oder Bakterien sein, welches die Endprodukte von Biosynthesewegen und Zwischenprodukte  
30 natürlich vorkommender Stoffwechselwege umfasst, sowie Moleküle, die im Stoffwechsel dieser Zellen nicht natürlich vorkommen, die jedoch von den erfindungsgemäßen Zellen produziert werden.

Eine weitere erfindungsgemäße Ausführungsform ist ein Verfahren  
35 zur Produktion von PUFAs, wobei das Verfahren das Züchten eines Organismus, der eine erfindungsgemäße Nukleinsäure, ein erfindungsgemäßes Genkonstrukt oder einen erfindungsgemäßen Vektor umfasst, welche ein Polypeptid kodieren, das C<sub>18</sub>-, C<sub>20</sub>- oder C<sub>22</sub>-Fettsäuren mit mindestens zwei Doppelbindungen im  
40 Fettsäuremolekül um mindestens zwei Kohlenstoffatome unter Bedingungen, unter denen PUFAs in dem Organismus produziert werden, verlängert, umfasst. Durch dieses Verfahren hergestellte PUFAs lassen sich durch Ernten der Organismen entweder aus der Kultur, in der sie wachsen, oder von dem Feld, Aufbrechen  
45 und/oder Extrahieren des geernteten Materials mit einem organischen Lösungsmittel isolieren. Aus diesem Lösungsmittel kann das Öl, das Lipide, Phospholipide, Sphingolipide, Glyco-

## 71

- lipide, Triacylglycerine und/oder freie Fettsäuren mit höherem Gehalt an PUFAs enthält, isoliert werden. Durch basische oder saure Hydrolyse der Lipide, Phospholipide, Sphingolipide, Glycolipide, Triacylglycerine können die freien Fettsäuren mit höherem
- 5 Gehalt an PUFAs isoliert werden. Ein höherer Gehalt an PUFAs bedeutet mindestens 5 %, vorzugsweise 10 %, besonders bevorzugt 20 %, ganz besonders bevorzugt 40 % mehr PUFAs als der ursprüngliche Organismus, der keine zusätzliche Nukleinsäure, die die erfindungsgemäße Desaturase kodiert, besitzt.
- 10 Vorzugsweise sind die durch dieses Verfahren produzierten PUFAs C<sub>18</sub>- oder C<sub>20-22</sub>-Fettsäuremoleküle mit mindestens zwei Doppelbindungen im Fettsäuremolekül, vorzugsweise drei, vier, bei Kombination mit einer weiteren Elongasen und einer  $\Delta$ -4 Desaturase
- 15 fünf oder sechs Doppelbindungen. Diese C<sub>18</sub>- oder C<sub>20-22</sub>-Fettsäuremoleküle lassen sich aus dem Organismus in Form eines Öls, Lipids oder einer freien Fettsäure isolieren. Geeignete Organismen sind beispielsweise die vorstehend erwähnten. Bevorzugte Organismen sind transgene Pflanzen.
- 20 Eine erfindungsgemäße Ausführungsform sind Öle, Lipide oder Fettsäuren oder Fraktionen davon, die durch das oben beschriebene Verfahren hergestellt worden sind, besonders bevorzugt Öl, Lipid oder eine Fettsäurezusammensetzung, die PUFAs umfassen und von
- 25 transgenen Pflanzen herrühren.
- Eine weitere erfindungsgemäße Ausführungsform ist die Verwendung des Öls, Lipids oder der Fettsäurezusammensetzung in Futtermitteln, Nahrungsmitteln, Kosmetika oder Pharmazeutika.
- 30 Ein weiterer Erfindungsgegenstand ist ein Verfahren zur Identifikation eines Antagonisten oder Agonisten von Desaturasen, umfassend
- 35 a) in Kontaktbringen der Zellen, die das Polypeptid der vorliegenden Erfindung exprimieren, mit einem Kandidatenstoff;
- b) Testen der Desaturaseaktivität;
- 40 c) Vergleichen der Desaturaseaktivität mit einer Standardaktivität in Abwesenheit des Kandidatenstoffs, wobei ein Anstieg der Desaturaseaktivität über den Standard anzeigt, daß der Kandidatenstoff ein Agonist und ein Verringerung
- 45 der Desaturaseaktivität anzeigt, daß der Kandidatenstoff ein Antagonist ist.

- Der genannte Kandidatenstoff kann ein chemisch synthetisierter oder mikrobiologisch produzierter Stoff sein und z.B. in Zell-extrakten von z.B. Pflanzen, Tieren oder Mikroorganismen auftreten. Weiterhin kann der genannte Stoff zwar im Stand der
- 5 Technik bekannt sein, aber bisher nicht bekannt sein als die Aktivität der Desaturasen steigernd oder reprimierend. Das Reaktionsgemisch kann ein zellfreier Extrakt sein oder eine Zelle oder Zellkultur umfassen. Geeignete Methoden sind dem Fachmann bekannt und werden z.B. allgemein beschrieben in Alberts, Molecular Biology the cell, 3<sup>rd</sup> Edition (1994), z.B. Kapitel 17. Die
- 10 genannten Stoffe können z.B. zu dem Reaktionsgemisch oder dem Kulturmedium zugegeben werden oder den Zellen injiziert werden oder auf eine Pflanze gesprüht werden.
- 15 Wenn eine Probe, die ein nach der erfindungsgemäßen Methode aktiven Stoff beinhaltet, identifiziert wurde, dann ist es entweder möglich, den Stoff direkt von der ursprünglichen Probe zu isolieren oder man kann die Probe in verschiedene Gruppen teilen, z.B. wenn sie aus einer Vielzahl von verschiedenen Komponenten
- 20 besteht, um so die Zahl der verschiedenen Substanzen pro Probe zu reduzieren und dann das erfindungsgemäße Verfahren mit einer solchen "Unterprobe" der ursprünglichen Probe zu wiederholen. Abhängig von der Komplexität der Probe können die oben beschriebenen Schritte mehrmals wiederholt werden, vorzugsweise bis die
- 25 gemäß der erfindungsgemäßen Methode identifizierte Probe nur noch eine geringe Anzahl von Substanzen oder nur noch eine Substanz umfaßt. Vorzugsweise wird der gemäß der erfindungsgemäßen Methode identifizierte Stoff oder Derivate davon weiter formuliert, so, daß er für die Anwendung in der Pflanzenzüchtung oder Pflanzen-
- 30 zell- oder Gewebekultur geeignet ist.

- Die Stoffe, die gemäß dem erfindungsgemäßen Verfahren getestet und identifiziert wurden, können sein: Expressionsbibliotheken, z.B. cDNA-Expressionsbibliotheken, Peptide, Proteine, Nukleinsäuren, Antikörper, kleine organische Stoffe, Hormone, PNAs oder
- 35 ähnliches (Milner, Nature Medicine 1 (1995), 879-880; Hupp, Cell. 83 (1995), 237-245; Gibbs, Cell. 79 (1994), 193-198 und darin zitierte Referenzen). Diese Stoffe könne auch funktionelle Derivate oder Analogon der bekannten Inhibitoren oder Aktivatoren
- 40 sein. Verfahren zur Herstellung von chemischen Derivaten oder Analogon sind dem Fachmann bekannt. Die genannten Derivate und Analogon können gemäß Verfahren nach dem Stand der Technik getestet werden. Weiterhin kann computergestütztes Design oder Peptidomimetics zur Herstellung geeigneter Derivate und Analogon
- 45 verwendet werden. Die Zelle oder das Gewebe, die/das für das erfindungsgemäße Verfahren verwendet werden kann, ist vorzugsweise eine erfindungsgemäße Wirtszelle, Pflanzenzelle oder ein

Pflanzengewebe, wie in den oben genannten Ausführungsformen beschrieben.

Entsprechend betrifft die vorliegende Erfindung auch einen  
5 Stoff, der gemäß den vorstehenden erfindungsgemäßen Verfahren  
identifiziert wurde. Der Stoff ist z.B. ein Homolog der  
erfindungsgemäßen Desaturasen. Homologe der Desaturasen können  
durch Mutagenese, z.B. durch Punktmutation oder Deletion der  
Desaturasen, erzeugt werden. Hierin verwendet wird der Begriff  
10 "Homolog" als eine variante Form der Desaturasen, die als Agonist  
oder Antagonist für die Aktivität der Desaturasen wirkt. Ein  
Agonist kann die im wesentlichen gleiche oder einen Teil der  
biologischen Aktivität der Desaturasen haben. Ein Antagonist der  
Desaturasen kann eine oder mehr Aktivitäten der natürlich vor-  
15 kommenden Formen der Desaturasen inhibieren, z.B. kompetitiv an  
ein *Downstream* oder *Upstream* gelegenes Mitglied der Fettsäure-  
synthese-Stoffwechselwege, die die Desaturasen einschließen,  
binden oder an Desaturasen binden und dabei die Aktivität  
reduzieren oder inhibieren.

20

Außerdem betrifft die vorliegende Erfindung auch ein Antikörper  
oder ein Fragment davon, wie sie hierin beschrieben werden, der  
die Aktivität der erfindungsgemäßen Desaturasen inhibiert.

25 Bei einem Aspekt betrifft die vorliegende Erfindung ein Anti-  
körper, der spezifisch den erfindungsgemäßen oben beschriebenen  
Agonisten oder Antagonisten erkennt bzw. bindet.

Ein weiterer Aspekt betrifft eine Zusammensetzung, die den Anti-  
30 körper, den nach dem erfindungsgemäßen Verfahren identifizierten  
Stopp oder das Antisense-Molekül umfaßt.

In einer weiteren Ausführungsform betrifft die vorliegende  
Erfindung ein Kit, umfassend die erfindungsgemäße Nukleinsäure,  
35 das erfindungsgemäße Genkonstrukt, die erfindungsgemäße Amino-  
säuresequenz, das erfindungsgemäße Antisense-Nukleinsäuremolekül,  
den erfindungsgemäßen Antikörper und/oder Zusammensetzung, einen  
Antagonisten oder Agonisten, der nach dem erfindungsgemäßen Ver-  
fahren hergestellt wurde, und/oder erfindungsgemäße Öle, Lipide  
40 und/oder Fettsäuren oder eine Fraktion davon. Ebenso kann das  
Kit die erfindungsgemäßen Wirtszellen, Organismen, Pflanzen  
oder Teile davon, erntbare Teile der erfindungsgemäßen Pflanzen  
oder Vermehrungsmaterial oder aber auch den erfindungsgemäßen  
Antagonisten oder Agonisten umfassen. Die Komponenten des Kits  
45 der vorliegenden Erfindung können in geeigneten Containern, bei-  
spielsweise mit oder in Puffern oder anderen Lösungen verpackt  
sein. Ein oder mehr der genannten Komponenten können in ein und

demselben Container verpackt sein. Zusätzlich oder alternativ können ein oder mehr der genannten Komponenten z.B. auf einer festen Oberfläche adsorbiert sein, z.B. Nitrozellulosefilter, Glasplatten, Chips, Nylonmembranen oder Mikrotiterplatten. Das  
5 Kit kann für jede der hierin beschriebenen Methoden und Ausführungsformen verwendet werden, z.B. für die Produktion von Wirtszellen, transgenen Pflanzen, zur Detektion von homologen Sequenzen, zur Identifikation von Antagonisten oder Agonisten usw. Weiterhin kann das Kit Anleitungen für die Verwendung des  
10 Kits für eine der genannten Anwendungen enthalten.

Diese Erfindung wird durch die nachstehenden Beispiele weiter veranschaulicht, die nicht als beschränkend aufgefaßt werden sollten. Der Inhalt sämtlicher in dieser Patentanmeldung  
15 zitierten Literaturstellen, Patentanmeldungen, Patente und veröffentlichten Patentanmeldungen ist hier durch Bezugnahme aufgenommen.

#### Beispielteil

20

#### Beispiel 1: Allgemeine Verfahren

##### a) Allgemeine Klonierungsverfahren:

25 Klonierungsverfahren, wie beispielsweise Restriktionsspaltungen, Agarosegelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrocellulose- und Nylonmembranen, Verbindung von DNA-Fragmenten, Transformation von Escherichia coli- und Hefe-Zellen, Anzucht von Bakterien und Sequenzanalyse  
30 rekombinanter DNA, wurden durchgeführt wie beschrieben in Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) oder Kaiser, Michaelis und Mitchell (1994) "Methods in Yeast Genetics" (Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3). Die Transformation und Anzucht von Algen,  
35 wie Chlorella oder Phaeodactylum werden durchgeführt wie beschrieben von El-Sheekh (1999), Biologia Plantarum 42:209-216; Apt et al. (1996) Molecular and General Genetics 252 (5):872-9.

##### b) Chemikalien

40

Die verwendeten Chemikalien wurden, wenn im Text nicht anders angegeben, in p. A.-Qualität von den Firmen Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen) bezogen. Lösungen wurden unter Verwendung  
45 von reinem pyrogenfreiem Wasser, im nachstehenden Text als H<sub>2</sub>O bezeichnet, aus einer Milli-Q-Wassersystem-Wasserreinigungsanlage (Millipore, Eschborn) hergestellt. Restriktionsendo-

## 75

nukleasen, DNA-modifizierende Enzyme und molekularbiologische Kits wurden bezogen von den Firmen AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/

5 Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) und Stratagene (Amsterdam, Niederlande). Wenn nicht anders angegeben, wurden sie nach den Anweisungen des Herstellers verwendet.

10

c) Zellmaterial

Die erfindungsgemäßen isolierten Nukleinsäuresequenzen sind im Genom eines *Phaeodactylum tricornutum* UTEX646-Stammes enthalten,  
15 der über die Algensammlung der University of Texas, Austin verfügbar ist.

*Phaeodactylum tricornutum* wurde bei 25°C mit einem Licht/Dunkel Rhythmus von 14:10 Stunden bei 22°C und 35 microEinstein (entspricht micromol Photonen pro Quadratmeter und Sekunde) in Glasröhren kultiviert, die von unten mit Luft begast wurden.  
20

Als Kulturmedium für *Phaeodactylum tricornutum* wurde das f/2 Kulturmedium mit 10 % organischen Medium nach Guillard, R.R.L. verwendet (1975; Culture of phytoplankton for feeding marine invertebrates. In: Smith, W.L. and Chanley, M.H. (Eds.) Culture of marine Invertebrate animals, NY Plenum Press, pp. 29-60.):  
25 Es enthält

30 995,5 ml Seewasser (artifizuell)  
1 ml NaNO<sub>3</sub> (75 g/l), 1 ml NaH<sub>2</sub>PO<sub>4</sub> (5 g/l), 1 ml Spurenelementelösung, 1 ml Tris/Cl pH 8.0, 0.5 ml f/2 Vitaminlösung

Spurenelementelösung: Na<sub>2</sub>EDTA (4,36 g/l), FeCl<sub>3</sub> (3,15 g/l),  
35 Primäre Spurenelemente: CuSO<sub>4</sub> (10 g/l), ZnSO<sub>4</sub> (22 g/l), CoCl<sub>2</sub> (10 g/l), MnCl<sub>2</sub> (18 g/l), NaMoO<sub>4</sub> (6,3 g/l)  
f/2 Vitaminlösung: Biotin: 10 mg/l, Thiamin 200 mg/l, Vit B12 0,1 mg/l

org-Medium: Na-Acetat (1 g/l), Glucose (6 g/l), Na-Succinat  
40 (3 g/l), Bacto-Trypton (4 g/l), Hefe-Extrakt (2 g/l)

45

## 76

Beispiel: 2 Isolierung von Gesamt-DNA aus *Phaeodactylum tricornutum* UTEX646 für Hybridisierungsexperimente

Die Einzelheiten der Isolierung von Gesamt-DNA betreffen die Auf-  
5 arbeitung von Pflanzenmaterial mit einem Frischgewicht von einem Gramm.

CTAB-Puffer: 2 % (Gew./Vol.) N-Acetyl-N,N,N-trimethylammonium-  
bromid (CTAB); 100 mM Tris-HCl, pH 8,0; 1,4 M NaCl; 20 mM EDTA.

10

N-Laurylsarkosin-Puffer: 10 % (Gew./Vol.) N-Laurylsarkosin;  
100 mM Tris-HCl, pH 8,0; 20 mM EDTA.

*Phaeodactylum tricornutum*-Zellmaterial wurde unter flüssigem  
15 Stickstoff in einem Mörser verrieben, so dass ein feines Pulver  
erhalten wurde, und in 2 ml-Eppendorfgefäße überführt. Das ge-  
frorene Pflanzenmaterial wurde dann mit einer Schicht von 1 ml  
Zersetzungspuffer (1 ml CTAB-Puffer, 100 ml N-Laurylsarkosin-  
Puffer, 20 ml  $\beta$ -Mercaptoethanol und 10 ml Proteinase K-Lösung,  
20 10 mg/ml) überschichtet und eine Stunde unter kontinuierlichem  
Schütteln bei 60°C inkubiert. Das erhaltene Homogenat wurde  
in zwei Eppendorfgefäße (2 ml) aufgeteilt und zweimal durch  
Schütteln mit dem gleichen Volumen Chloroform/Isoamylalkohol  
(24:1) extrahiert. Zur Phasentrennung wurde eine Zentrifugation  
25 bei 8000 x g und RT (= Raumtemperatur = ~ 23°C) jeweils 15 min  
lang durchgeführt. Die DNA wurde dann 30 min unter Verwendung von  
eiskaltem Isopropanol bei -70°C gefällt. Die gefällte DNA wurde  
bei 10000 g 30 min bei 4°C sedimentiert und in 180 ml TE-Puffer  
(Sambrook et al., 1989, Cold Spring Harbor Laboratory Press: ISBN  
30 0-87969-309-6) resuspendiert. Zur weiteren Reinigung wurde die  
DNA mit NaCl (1,2 M Endkonzentration) behandelt und erneut 30 min  
unter Verwendung des zweifachen Volumens an absolutem Ethanol bei  
-70°C gefällt. Nach einem Waschschrift mit 70 % Ethanol wurde die  
DNA getrocknet und anschließend in 50 ml H<sub>2</sub>O + RNase (50 mg/ml  
35 Endkonzentration) aufgenommen. Die DNA wurde über Nacht bei 4°C  
gelöst und die RNase-Spaltung wurde anschließend 1 Std. bei  
37°C durchgeführt. Die Aufbewahrung der DNA erfolgte bei 4°C.

Beispiel 3: Isolierung von Gesamt-RNA und poly(A)<sup>+</sup>-RNA aus  
40 Pflanzen und *Phaeodactylum tricornutum*

Die Isolierung von Gesamt-RNA aus Pflanzen wie Lein und Raps  
etc. erfolgt nach einer bei Logemann et al beschriebenen  
Methode (1987, Anal. Biochem. 163, 21) isoliert. Aus Moos kann  
45 die Gesamt-RNA *Protonema*-Gewebe nach dem GTC-Verfahren (Reski  
et al., 1994, Mol. Gen. Genet., 244:352-359) gewonnen werden.

RNA Isolierung aus *Phaeodactylum tricornutum*:

Tiefgefrorene Algenproben (-70°C) wurden in einem eiskaltem Mörser unter Flüssigstickstoff zu feinem Pulver zerreiben.

- 5 2 Volumen Homogenisationsmedium (12,024 g Sorbitol, 40,0 ml 1M Tris-HCl, pH 9 (0,2 M); 12,0 ml 5 M NaCl (0,3 M), 8,0 ml 250 mM EDTA, 761,0 mg EGTA, 40,0 ml 10 % SDS wurden auf 200 ml mit H<sub>2</sub>O aufgefüllt und der pH auf 8,5 eingestellt) und 4 Volumen Phenol mit 0,2 % Mercaptoethanol wurden bei 40 bis 50°C unter gutem
- 10 Mischen zu gefrorenem Zellpulver gegeben. Danach wurden 2 Volumen Chloroform hinzufügen und für 15 min kräftig gerührt. Es wurde 10 min bei 10000 g zentrifugiert und die wässrige Phase mit Phenol/Chloroform (2 Vol) und abschließend mit Chloroform extrahiert.

15

Das erhaltene Volumen der wässrigen Phase wurde mit 1/20 Vol 4 M Na-Acetat (pH 6) und 1 Vol Isopropanol (eiskalt) versetzt und die Nukleinsäuren bei -20°C über Nacht (= ÜN) gefällt. Anschließend wurde 30 min bei 10000 g zentrifugiert und der

- 20 Überstand abgesogen. Es folgte ein Waschschrift mit 70 % EtOH und erneute Zentrifugation. Das Sediment wurde in Tris-Borat-Puffer (80 mM Tris-Borat-Puffer, 10 mM EDTA, pH 7,0) aufgenommen. Dann wurde der Überstand mit 1/3 Vol 8 M LiCl versetzt, gemischt 30 min bei 4°C inkubiert. Nach erneutem zentrifugieren wurde
- 25 das Sediment mit 70 % Ethanol gewaschen, zentrifugiert und das Sediment in RNase-freiem Wasser gelöst.

Die Isolierung von poly(A)<sup>+</sup>-RNA erfolgte unter Verwendung von Dyna Beads® (Dyna, Oslo, Finnland) nach den Anweisungen im Protokoll

- 30 des Herstellers.

Nach der Bestimmung der RNA- oder poly(A)<sup>+</sup>-RNA-Konzentration wurde die RNA durch Zugabe von 1/10 Volumina 3 M Natriumacetat, pH 4,6, und 2 Volumina Ethanol gefällt und bei -70°C aufbewahrt.

35

Für die Analyse wurden jeweils 20 µg RNA in einem Formaldehyd-haltigen 1,5%igen Agarosegel aufgetrennt und auf Nylon Membranen (Hybond, Amersham) überführt. Der Nachweis spezifischer Transkripte wurde wie bei Amasino beschrieben durchgeführt

- 40 ((1986) Anal. Biochem. 152, 304)).

## Beispiel 4: Konstruktion der cDNA-Bank

Zur Konstruktion der cDNA-Bank aus *Phaeodactylum tricornutum*

- 45 wurde die Erststrangsynthese unter Verwendung von Reverser Transkriptase aus Maus-Leukämie-Virus (Roche, Mannheim, Deutschland) und Oligo-d(T)-Primern, die Zweitstrangsynthese

durch Inkubation mit DNA-Polymerase I, Klenow-Enzym und RNase H-Spaltung bei 12°C (2 Std.), 16°C (1 Std.) und 22°C (1 Std.) erzielt. Die Reaktion wurde durch Inkubation bei 65°C (10 min) gestoppt und anschließend auf Eis überführt. Doppelsträngige DNA-Moleküle wurde mit T4-DNA-Polymerase (Roche, Mannheim) bei 37°C (30 min) mit glatten Enden versehen. Die Nukleotide wurden durch Phenol/Chloroform-Extraktion und Sephadex-G50-Zentrifugiersäulen entfernt. EcoRI/XhoI-Adapter (Pharmacia, Freiburg, Deutschland) wurden mittels T4-DNA-Ligase (Roche, 12°C, über Nacht) an die cDNA-Enden ligiert, mit XhoI nachgeschnitten und durch Inkubation mit Polynukleotidkinase (Roche, 37°C, 30 min) phosphoryliert. Dieses Gemisch wurde der Trennung auf einem Low-Melting-Agarose-Gel unterworfen. DNA-Moleküle über 300 Basenpaaren wurden aus dem Gel eluiert, Phenol-extrahiert, auf Elutip-D-Säulen (Schleicher und Schüll, Dassel, Deutschland) konzentriert und an Vektorarme ligiert und in lambda-ZAP-Express-Phagen unter Verwendung des Gigapack Gold-Kits (Stratagene, Amsterdam, Niederlande) verpackt, wobei Material des Herstellers verwendet und seine Anweisungen befolgt wurden.

20

#### Beispiel 5: DNA-Sequenzierung und Computeranalyse

cDNA-Banken, wie im Beispiel 4 beschrieben, wurden zur DNA-Sequenzierung nach Standardverfahren, insbesondere durch das Kettenterminationsverfahren unter Verwendung des ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction-Kit (Perkin-Elmer, Weiterstadt, Deutschland), verwendet. Die Sequenzierung zufälliger, vereinzelter Klone wurde anschließend an die präparative Plasmidgewinnung aus cDNA-Banken über in vivo-Massen-excision und Retransformation von DH10B auf Agarplatten durchgeführt (Einzelheiten zu Material und Protokoll von Stratagene, Amsterdam, Niederlande). Plasmid-DNA wurde aus über Nacht gezüchteten E. coli-Kulturen, die in Luria-Brühe mit Ampicillin (siehe Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)) gezüchtet worden waren, an einem Qiagen-DNA-Präparations-Roboter (Qiagen, Hilden) nach den Protokollen des Herstellers präpariert. Sequenzierprimer mit den folgenden Nukleotidsequenzen wurden verwendet:

40 5'-CAGGAAACAGCTATGACC-3'  
 5'-CTAAAGGGAACAAAAGCTG-3'  
 5'-TGTAACACGACGGCCAGT-3'

Die Sequenzen wurden unter Verwendung des Standard-Softwarepakets EST-MAX, das kommerziell von Bio-Max (München, Deutschland) geliefert wird, prozessiert und annotiert. Durch Nutzung von Vergleichsalgorithmen und unter Verwendung der in SEQ ID NO: 8

dargestellten Suchsequenz wurde mithilfe des BLAST-Programms nach homologen Genen gesucht (Altschul et al. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.). Zwei Sequenzen aus *Phaeodactylum tricornutum* mit Homologien zur Suchsequenz aus *Physcomitrella patens* wurden eingehender charakterisiert.

Beispiel 5a: Isolation von Desaturasen aus *Phaeodactylum tricornutum* über Polymerase Kettenreaktion mithilfe degenerierter Oligonukleotide:

Mithilfe von publizierten Desaturasen können Motive identifiziert werden, die für  $\Delta$ -5 und  $\Delta$ -6 Desaturasen typisch sind. Im folgenden sind Oligonukleotidsequenzen mit möglichen Variationen dargestellt. Unter der Oligonukleotidsequenz ist im Ein-Buchstaben-code die Aminosäure dargestellt, von der die Basenkombination abgeleitet werden kann. Z.B. bedeutet A/G, daß an dieser Position bei der Synthese des Bausteins statistisch gleichverteilt entweder ein A oder ein G in das Oligonukleotid eingebaut wird, da das von der korrespondierenden Aminosäure abgeleitete Basentriplett entweder ein AAA oder ein AAG sein kann. Die DNA Sequenz kann auch ein Inosin (i) enthalten, wenn die Bestimmung einer Base an dieser Position aufgrund des genetischen Codes drei oder vier unterschiedliche Basen erlaubt. Folgende Sequenzen und Primer können verwendet werden:

5'-Vorwärts-Primer:

F1a:	TGG	TGG	AA	A/G	TGG	AAi	CA	T/C	AA				
F1b:	TGG	TGG	AA	A/G	TGG	ACi	CA	T/C	AA				
30 F1a:	W	W	K		W	N/T	H		K/N				
F1b:	W	W	K		W	K	H		K/N				
F2a:	Gi	TGG	AA	A/G	GAi	A/C	Ai	CA	T/C	AA			
F2b:	Gi	TGG	AA	A/G	TTG	A/C	Ai	CA	T/C	AA			
35 F2a:	G/W	W	K		E/D	K/Q/N	H			K/N			
F2b:	G/W	W	K		W	K/Q/N	H			K/N			
F3a:	T	A/T	i		TTG	AAi	A/C	A	A/G	C/A	G/A	i	CA
F3b:	T	A/T	i		TTG	AAi	A/C	A	A/G	CAi			CA
F3a:	W				W	K/N	H/N			R/Q			H
40 F3b:	Y				W	K/N	H/N			R/Q			H
F4a:		GTi	TGG	A	A/T	G/A	GA	A/G		CA	A/G	CA	
F4b:		GTi	TGG	A	A/T	G/A	A/T	A	T/C	CA	A/G	CA	
F4a:		V	W		K/M		E			Q			H
F4b:		V	W		K/M		N/Y			Q			H
45 F5a1:		CA	T/C	TA	T/C	TGG	AA	A/G	AA	T/C	CA	G	C
F5a1:		CA	T/C	TA	T/C	TGG	AA	A/G	AA	T/C	CA	A	C
F5a1:		H		Y		W	K		N		Q		H/Q

## 80

F6a: TTG TTG AAi A/C A A/G AA i CA T/C AA  
 F6a: W W K/N H/N K/N H K/N

## 3'- Reverse Primer

5 R1b: GG A/G AA iAG G/A TG G/A TG T/C TC  
 R1b: GG A/G AA iAA G/A TG G/A TG T/C TC  
 R1a: P F L H H E  
 R1b: P F F H H E  
 R2a1: AA iAG A/G TG A/G TG iA C/T iA/G T/C TG  
 10 R2a2: AA T/C AA A/G TG A/G TG iA C/T iA/G T/C TG  
 R2a1: F L H H V/I V/A Q  
 R3a1: AT iTG iGG A/G AA iAA A/G TG A/G TG  
 R3a2: AT A/G TT iGG A/G AA iAA A/G TG A/G TG  
 R3a3: AT iTG iGG A/G AA iAG A/G TG A/G TG  
 15 R3a4: AT A/G TT iGG A/G AA iAG A/G TG A/G TG  
 R3a1: I/M H/Q P F F H H  
 R3a2: I/M N P F L H H  
 R4a1: CT iGG A/G AA iA A/G A/G TG A/G TG  
 R4a2: GA iGG A/G AA iA A/G A/G TG A/G TG  
 20 R4a3: GT iGG A/G AA iA A/G A/G TG A/G TG  
 R4a1: = T/R/S P F F/L H H  
 R5a1: AA iAA A/G TG A/G TG T/C TC T/A/G AT T/C TG  
 R5a2: AA iAG A/G TG A/G TG T/C TC T/A/G AT T/C TG  
 R5a1: F F H H E I Q  
 25 R5a2: F L H H E I Q  
 R6a1: T iGG iA A/G iAA A/G TG A/G TG iAC  
 R6a1: T iGG iA A/G iAG A/G TG A/G TG iAC  
 R6a1: T/N P L F/L H H V

- 30 Aufgrund verschiedener Variationsmöglichkeiten sind viele abgeleitete Oligonukleotide möglich, jedoch überraschenderweise gefunden wurde, dass dargestellte Oligonukleotide besonders zur Isolation von Desaturasen geeignet sein können.
- 35 Die Primer können in allen Kombinationen für Polymerase Kettenreaktionen eingesetzt werden. Mithilfe einzelner Kombinationen konnten Desaturase-Fragmente isoliert, wenn nachfolgende Bedingungen berücksichtigt wurden: Für PCR Reaktionen wurden jeweils 10 nMol Primer und 10 ng einer durch in vivo Excision
- 40 gewonnenen Plasmidbank eingesetzt. Die Plasmidbank konnte nach Protokollen des Herstellers (Stratagene) aus der Phagenbank isoliert werden. Die PCR-Reaktion wurde in einem Thermocycler (Biometra) mit der Pfu-DNA-Polymerase (Stratagene) und dem folgenden Temperaturprogramm durchgeführt: 3 min bei 96°C, gefolgt
- 45 von 35 Zyklen mit 30 s bei 96°C, 30 s bei 55°C und 1 min bei 72°C. Dabei wurde die Anlagerungstemperatur nach dem ersten Schritt von 55°C schrittweise um je 3°C erniedrigt und nach dem fünften Zyklus

eine Anlagerungstemperatur von 40°C beibehalten. Letztlich wurde ein Zyklus mit 10 min bei 72°C durchgeführt und der Ansatz durch Kühlen auf 4°C beendet.

- 5 Die Primerkombination F6a und R4a2 sind im Text unterstrichen gekennzeichnet und konnten erfolgreich zur Isolierung eines Desaturasefragmentes genutzt werden. Das resultierende Fragment konnte durch Sequenzierung verifiziert werden und zeigte Homologien zu einer Desaturase mit der Genbank Accession Nr. T36617
- 10 aus *Streptomyces coelicolor*. Die Homologie wurde mithilfe des BLASTP Programmes erhalten. Der Vergleich ist in Figur 4 dargestellt. Es ergaben sich Identitäten von 34 % und eine Homologie von 43 % zu Sequenz T36617. Das DNA-Fragment wurde gemäß Beispiel 7 in einem Hybridisierungsexperiment zur Isolierung
- 15 eines Volllängengens nach Standardbedingungen erfindungsgemäß eingesetzt.

- Die Codierregion einer so isolierten DNA-Sequenz wurde durch Übersetzung des genetischen Codes in eine Polypeptidsequenz
- 20 erhalten. In SEQ ID NO: 3 ist eine 1434 Basenpaare lange Sequenz dargestellt, die durch beschriebenes Verfahren isoliert werden konnte. Die Sequenz besitzt ein Startcodon in Position 1 bis 3 und ein Stopcodon in Position 1432-1434 und konnte in ein 477 Aminosäuren langes Polypeptid übersetzt werden. Durch Vergleich
- 25 gleich mit einer in WO 98 46763 beschriebenen Gensequenz wurde gefunden, dass ein nicht identisches aber homologes Fragment aus *Phaeodactylum tricornutum* codierend für 87 Aminosäuren vorbeschrieben wurde. Jedoch offenbart WO 98/46763 weder eine vollständige, funktionell aktive Desaturase noch Positions-
- 30 oder Substratspezifität. Dies wird auch dadurch deutlich, dass sowohl Homologien zur  $\Delta$ -5, als auch zur  $\Delta$ -6-Desaturase aus *Mortierella alpina* berichtet werden, ohne eine genaue Funktion festzulegen. Die erfindungsgemäße Sequenz hingegen codiert für eine funktionell aktive  $\Delta$ -6-Acyl Lipid Desaturase.

35

Beispiel 6: Identifizierung von DNA Sequenzen codierend für Desaturasen aus *Phaeodactylum tricornutum*

- Die Volllängensequenz der  $\Delta$ -6-Acyl Lipid Desaturase Pp\_des6
- 40 AJ222980 (NCBI Genbank Accession Nr.) aus dem Moos *Physcomitrella patens* (siehe auch Tabelle 1) sowie die  $\Delta$ -12-acyl Lipid Desaturase Sequenz (Tabelle 1 siehe Ma\_des12) aus *Mortierella alpina* AF110509 (AF110509 NCBI Genbank Accession Nr.) wurden für Sequenzvergleiche mithilfe des TBLASTN Suchalgorithmus
- 45 eingesetzt.

Die EST-Sequenzen PT0010070010R, PT001072031R sowie PT001078032R wurden zunächst aufgrund schwacher Homologien mit den Suchsequenzen aus *Physcomitrella* und *Mortierella* unter weiteren Kandidatengen als Zielgen in Betracht gezogen. In Figur 1 und 5 in Figur 2 sowie Figur 2a ist das Ergebnis der zwei gefundenen est-Sequenzen dargestellt. Die gefundenen Sequenzen sind Teil der erfindungsgemäßen Nukleinsäuren aus SEQ ID NO: 1 (Genname: Pt\_des5, eigene Datenbank Nr. der Erfinder PT001078032R), SEQ ID NO: 5. (Genname: Pt\_des12, eigene Datenbank NR. der 10 Erfinder PT0010070010R) und SEQ ID NO: 11 (Genname: Pt\_des12.2, eigene Datenbank des Erfinders PT001072031R). Buchstaben zeigen identische Aminosäuren an, während das Pluszeichen eine chemisch ähnliche Aminosäure bedeutet. Die Identitäten bzw. Homologien aller erfindungsgemäß gefundener Sequenzen gehen aus Tabelle 2 15 zusammenfassend hervor.

Desaturasen können Cytochrom b5 Domänen aufweisen, die auch in anderen nicht Desaturasen codierenden Genen vorkommen. Cytochrom b5 Domänen zeigen mithin hohe Homologien an, obwohl es sich um 20 verschiedene Genfunktionen handelt. Desaturasen können schwach konservierter Bereiche lediglich als putative Kandidatengene identifiziert werden und müssen auf die Enzymaktivität und Positionsspezifität der enzymatischen Funktion hin geprüft werden. Beispielsweise zeigen auch verschiedene Hydroxylasen, 25 Acetylasen und Epoxygenasen ähnlich wie Desaturasen Histidin-Box Motive, so dass eine konkrete Funktion experimentell nachgewiesen werden muß und zusätzlich die Verifizierung der Doppelbindung erst eine sichere Enzymaktivität und Positionsspezifität einer Desaturase ermöglicht. Überraschenderweise wurde gefunden, 30 dass erfindungsgemäße  $\Delta$ -6- und  $\Delta$ -5- Desaturase besonders geeignete Substratspezifitäten aufweisen und besonders geeignet sind, um in Kombination mit einer  $\Delta$ -6-Elongase aus *Physcomitrella* zur Produktion von polyungesättigten Fettsäuren wie Arachidonsäure, Eicosapentaensäure und Docosahexaensäure genutzt werden können.

35 Die Sequenzierung des vollständigen cDNA Fragmentes aus Klon PT001078032R ergab eine 1652 Basenpaare lange Sequenz. Die Sequenz codiert für ein Polypeptid von 469 Aminosäuren dargestellt in SEQ ID NO: 2. Diese wurde erhalten durch Übersetzung des genetischen Codes aus SEQ ID NO: 1 mit einem Startcodon in Basenpaarposition 115-117 und mit einem Stopcodon in Basenpaarposition 1522-1524. Der Klon beinhaltet ein vollständiges Desaturase-Polypeptid, wie aus dem Sequenzvergleich in Figur 3 zu ersehen ist. Striche bedeuten identische Amino- 45 säuren während Doppelpunkte und Einzelpunkte chemisch austauschbare, d.h. chemisch äquivalente Aminosäuren darstellen. Der Vergleich wurde mit der BLOSUM62 Austauschmatrix für Amino-

säuren nach Henikoff & Henikoff durchgeführt: ((1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919). Verwendete Parameter: Gap Weight: 8; Average Match: 2.912, Length Weight: 2, Average Mismatch: -2.003.

5

In Figur 6 und Figur 7 ist der Vergleich der MA\_des12 Peptidsequenz mit den gefundenen Sequenzen dargestellt.

Die Sequenzierung des vollständigen cDNA Fragmentes aus Klon  
10 PT0010070010R ergab eine in SEQ ID NO: 5 dargestellte 1651 Basenpaare lange Sequenz mit einem Startcodon in Position 67-69 und einem Stopcodon in Position 1552-1554. Die erfindungsgemäße Polypeptidsequenz ist in SEQ ID NO: 6 dargestellt.

15 Die Sequenzierung des vollständigen identifizierten cDNA Fragmentes aus Klon PT0010072031R ergab eine in SEQ ID NO: 11 dargestellte 1526 Basenpaare lange Sequenz mit einem Startcodon in Position 92-94 und einem Stopcodon in Position 1400-1402. Die erfindungsgemäße Polypeptidsequenz ist in SEQ ID NO: 12  
20 dargestellt.

In Tabelle 2 sind die Identitäten und Homologien erfindungsgemäßer Desaturasen untereinander und mit der Desaturase aus *Physcomitrella patens* und *Mortierella alpina* dargestellt. Die  
25 Angaben wurden mithilfe des Programms Bestfit unter gegebenen Parametern wie unten definiert als Teilprogramm folgender Software erhalten: Wisconsin Package Version 10.0 (Genetics Computer Group (GCG), Madison, Wisc., USA). Henikoff, S. and Henikoff, J.G. (1992). Amino acid substitution matrices from protein  
30 blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.

Weiterhin ist in Figur 5 der Vergleich der  $\Delta$ -6-acyl Lipid Desaturase aus *Physcomitrella patens* mit der Polypeptidsequenz des Klons Pt\_des6 dargestellt.

35

Tabelle 2:

	Homologie /	Suchsequenz	Suchsequenz
	Identität in %	Pp_des6	Ma_des12
40	Pt_des5	34.92/26.37	n.d.
	Pt_des6	50.69/41.06	n.d.
	Pt_des12	n.d.	48.58/38.92
	Pt_des12.2	n.d.	48.37/41.60

45 n.d. = nicht durchgeführt

Mithilfe des Algorithmus TBLASTN 2.0.10: Altschul et al 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402 wurden über einen lokalen Datenbankvergleich Sequenzen mit höchster Sequenz-  
 5 homologie bzw. Identität identifiziert. Die Ergebnisse sind in folgender Tabelle 2A dargestellt.

Tabelle 2A: Homologe mit den höchsten Sequenzhomologien bzw  
 Identitäten zu erfindungsgemäßen Polypeptidsequenzen  
 10 aus SEQ ID NO. 2, 4, 6 oder 12

Homologie / Identität(%)	Suchsequenz PT001070010R	Suchsequenz PT001072031R	Suchsequenz PT001078032R	Suchsequenz Pt_des6
15 L26296: Fad2 A. thaliana	50 % / 37 %	n.d.	n.d.	n.d.
U86072 Petro- selinum crispum Fad2	n.d.	51/40	n.d.	n.d.
20 AL358652 L. major putative desaturase	n.d.	n.d.	45/30	n.d.
25 AB020032 M. alpina delta 6 desaturase	n.d.	n.d.	n.d.	53/38

#### Beispiel 7: Identifikation von Genen mittels Hybridisierung

30 Gensequenzen lassen sich zur Identifikation homologer oder heterologer Gene aus cDNA- oder genomischen Banken verwenden.

Homologe Gene (d.h. Voll-Längen-cDNA-Klone, die homolog sind, oder Homologen) lassen sich über Nukleinsäurehybridisierung unter Verwendung von beispielsweise cDNA-Banken isolieren: Ins-  
 35 besondere zur Isolierung von funktionell aktiven Voll-Längengenen der in SEQ ID NO: 3 gezeigten kann die Methode genutzt werden. Je nach der Häufigkeit des Gens von Interesse werden 100000 bis zu 1000000 rekombinante Bakteriophagen plattiert und auf eine  
 40 Nylonmembran überführt. Nach der Denaturierung mit Alkali wurde die DNA auf der Membran z.B. durch UV-Vernetzung immobilisiert. Die Hybridisierung erfolgt bei hoch-stringenten Bedingungen. In wässriger Lösung werden die Hybridisierung und die Waschschr  
 45 mittels radioaktiver (<sup>32</sup>P-) Nicktranskription (High Prime, Roche, ..

Mannheim, Deutschland) hergestellt. Die Signale werden mittels Autoradiographie nachgewiesen.

- Partiell homologe oder heterologe Gene, die verwandt, aber nicht  
5 identisch sind, lassen sich analog zum oben beschriebenen Verfahren unter Verwendung niedrig-stringenter Hybridisierungs- und Waschbedingungen identifizieren. Für die wässrige Hybridisierung wurde die Ionenstärke gewöhnlich bei 1 M NaCl gehalten, wobei die Temperatur nach und nach von 68 auf 42°C gesenkt wurde.
- 10 Die Isolierung von Gensequenzen, die nur zu einer einzelnen Domäne von beispielsweise 10 bis 20 Aminosäuren Homologien aufweisen, lässt sich unter Verwendung synthetischer, radioaktiv markierter Oligonukleotidsonden durchführen. Radioaktiv markierte  
15 Oligonukleotide werden mittels Phosphorylierung des 5'-Endes zweier komplementärer Oligonukleotide mit T4-Polynukleotidkinase hergestellt. Die komplementären Oligonukleotide werden aneinander hybridisiert und ligiert, so dass Konkatemere entstehen. Die doppelsträngigen Konkatemere werden beispielsweise durch Nick-  
20 transkription radioaktiv markiert. Die Hybridisierung erfolgt gewöhnlich bei niedrig-stringenten Bedingungen unter Verwendung hoher Oligonukleotidkonzentrationen.

Oligonukleotid-Hybridisierungslösung:

- 25 6 x SSC  
0,01 M Natriumphosphat  
1 mM EDTA (pH 8)  
0,5 % SDS
- 30 100 mikrog/ml denaturierte Lachssperma-DNA  
0,1 % fettarme Trockenmilch

- Während der Hybridisierung wird die Temperatur schrittweise auf 5 bis 10°C unter die berechnete Oligonukleotid-T<sub>m</sub> oder bis auf Raum-  
35 temperatur (bedeutet RT = ~ 23°C in allen Experimenten, wenn nicht anders angegeben) gesenkt, gefolgt von Waschschritten und Autoradiographie. Das Waschen wird mit extrem niedriger Stringenz durchgeführt, zum Beispiel 3 Waschschriffe unter Verwendung von 4 X SSC. Weitere Einzelheiten sind wie von Sambrook, J., et al.  
40 (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, oder Ausubel, F.M., et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons, beschrieben.

Beispiel 8: Identifikation von Zielgenen durch Sichtung von  
Expressionsbanken mit Antikörpern

Es wurden cDNA-Sequenzen zur Herstellung von rekombinanten  
5 Protein zum Beispiel in *E. coli* verwendet (z.B. Qiagen QIAexpress  
pQE-System). Die rekombinanten Proteine wurden dann gewöhnlich  
über Ni-NTA-Affinitätschromatographie (Qiagen) affinitäts-  
gereinigt. Die rekombinanten Proteine wurden dann zur Herstellung  
spezifischer Antikörper beispielsweise unter Verwendung von  
10 Standardtechniken zur Immunisierung von Kaninchen verwendet.  
Anschließend wurden die Antikörper dann unter Verwendung einer  
Ni-NTA-Säule, die mit rekombinantem Antigen vorgesättigt wird,  
affinitätsgereinigt, wie von Gu et al., (1994) *BioTechniques*  
17:257-262 beschrieben. Der Antikörper kann dann zur Durch-  
15 musterung von Expressions-cDNA-Banken mittels immunologischem  
Sichtung verwendet werden (Sambrook, J., et al. (1989),  
"Molecular Cloning: A Laboratory Manual", Cold Spring Harbor  
Laboratory Press, oder Ausubel, F.M., et al. (1994) "Current  
Protocols in Molecular Biology", John Wiley & Sons).

20

Beispiel 9: Transformation von *Agrobacterium*

Die *Agrobacterium*-vermittelte Pflanzentransformation kann zum  
Beispiel unter Verwendung des GV3101- (pMP90-) (Koncz und Schell,  
25 *Mol. Gen. Genet.* 204 (1986) 383-396) oder LBA4404- (Clontech)  
oder C58C1 pGV2260 (Deblaere et al 1984, *Nucl. Acids Res.* 13,  
4777-4788) *Agrobacterium tumefaciens*-Stamms durchgeführt werden.  
Die Transformation kann durch Standard-Transformationstechniken  
durchgeführt werden (ebenfalls Deblaere et al. 1984).

30

Beispiel 10: Pflanzentransformation

Die *Agrobacterium*-vermittelte Pflanzentransformation kann unter  
Verwendung von Standard-Transformations- und Regenerations-  
35 techniken durchgeführt werden (Gelvin, Stanton B., Schilperoort,  
Robert A., *Plant Molecular Biology Manual*, 2. Aufl., Dordrecht:  
Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale Signatur:  
BT11-P ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E.,  
*Methods in Plant Molecular Biology and Biotechnology*, Boca Raton:  
40 CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

Beispielsweise kann Raps mittels Kotyledonen- oder Hypokotyl-  
transformation transformiert werden (Moloney et al., *Plant*  
*Cell* 8 (1989) 238-242; De Block et al., *Plant Physiol.* 91 (1989)  
45 694-701). Die Verwendung von Antibiotika für die *Agrobacterium*-  
und Pflanzenselektion hängt von dem für die Transformation  
verwendeten binären Vektor und *Agrobacterium*-Stamm ab. Die

Rapsselektion wird gewöhnlich unter Verwendung von Kanamycin als selektierbarem Pflanzenmarker durchgeführt.

Der Agrobacterium-vermittelte Gentransfer in Lein (*Linum usitatissimum*) lässt sich unter Verwendung von beispielsweise einer von Mlynarova et al. (1994) Plant Cell Report 13:282-285 beschriebenen Technik durchführen.

Die Transformation von Soja kann unter Verwendung von beispielsweise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International) oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University Toledo) beschriebenen Technik durchgeführt werden.

Die Pflanzentransformation unter Verwendung von Teilchenbeschuss, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben von Freeling und Walbot "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York).

#### 20 Beispiel 11: Plasmide für die Pflanzentransformation

Geeignete binäre Vektoren und Transformationsmarker

Zur Pflanzentransformation können binäre Vektoren, wie pBinAR (Höfgen und Willmitzer, Plant Science 66 (1990) 221-230) oder pGPTV (Becker et al. 1992, Plant Mol. Biol. 20:1195-1197) verwendet werden. Die Konstruktion der binären Vektoren kann durch Ligation der cDNA in Sense- oder Antisense-Orientierung in T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3' von der cDNA. Die binären Vektoren können unterschiedliche Markergene tragen. So kann etwa die Resistenz durch die Expression des nptII Markergens unter Kontrolle des 35S oder des nos Promoters erfolgen. Insbesondere kann das nptII-Markergen codierend für Kanamycin-Resistenz vermittelt durch Neomycinphosphotransferase gegen die herbizidresistente Form eines Acetolactat Synthasegens (AHAS oder ALS) ausgetauscht werden. Das ALS-Gen ist beschrieben in Ott et al., J. Mol. Biol. 1996, 263:359-360. Geeignet für manche Pflanzen ist auch die Verwendung des Hygromycin-Resistenz-Gens. Der v-ATPase-cl-Promotor kann in das Plasmid pBin19 oder pGPTV kloniert werden und durch Klonierung vor die kodierende Region des ALS Gens für die Markergenexpression genutzt werden. Der genannte v-ATPase-cl-Promotor entspricht einem 1153 Basenpaarfragment aus Beta vulgaris (Plant Mol Biol, 1999, 39:463-475). Der genannte nos Promoter Dabei können sowohl Sulphonylharnstoffe als auch Imidazolinone wie Imazethapyr oder Sulphonylharnstoffe als Antimetaboliten zur Selektion verwendet wer-

den. Alternativ kann auch der nos-Promoter für die Markergenexpression verwendet werden.

Beispiel 12: Ermittlung von geeigneten Promotoren für die  
Expression in Lein

Die gewebespezifische Expression lässt sich unter Verwendung eines gewebespezifischen Promotors erzielen. Beispielsweise kann die samenspezifische Expression erreicht werden, indem der DC3- oder der LeB4- oder der USP-Promotor oder der Phaseolin-Promotor 5' der cDNA einkloniert wird. Auch jedes andere samenspezifische Promotorelement wie z.B. der Napin- oder Arcelin Promotor (Goossens et al. 1999, Plant Phys. 120(4):1095-1103 und Gerhardt et al. 2000, Biochimica et Biophysica Acta 1490(1-2):87-98) kann verwendet werden. Zur konstitutiven Expression in der ganzen Pflanze lässt sich der CaMV-35S-Promotor oder ein v-ATPase-c1 Promotor verwenden.

Um die Eigenschaften des Promotors zu bestimmen und die essentiellen Elemente desselben, die seine Gewebespezifität ausmachen, zu identifizieren, ist es erforderlich, den Promotor selbst oder verschiedene Fragmente desselben vor ein sogenanntes Reportergen zu setzen, das eine Bestimmung der Expressionsaktivität ermöglicht. Beispielhaft für ein Reportergen sei die bakterielle  $\beta$ -Glucuronidase (GUS) genannt (Jefferson et al., EMBO J. 1987, 6, 3901-3907). Die  $\beta$ -Glucuronidase Aktivität kann *in-situ* mittels eines chromogenen Substrates wie 5-Bromo-4-Chloro-3-Indolyl- $\beta$ -D-Glucuronsäure im Rahmen einer Aktivitätsfärbung bestimmt werden (Jefferson, 1987, Plant Molecular Biology Reporter 5, 387-405). Für die Untersuchungen der Gewebespezifität wird das pflanzliche Gewebe geschnitten, eingebettet, gefärbt und analysiert wie beschrieben (z.B. Bäumlein H et al., 1991 Mol Gen Genet 225: 121-128).

Fluorimetrischer GUS-Test (nach Montgomery et al., 1993)

Dieser Assay erlaubt eine quantitative Bestimmung der GUS-Aktivität in dem untersuchten Gewebe. Für die quantitative Aktivitätsbestimmung wird als Substrat für die  $\beta$ -Glucuronidase MUG (4-Methyl-umbelliferyl-beta-D-glucuronid) verwendet, das in MU (Methyl-umbelliferon) und Glucuronsäure gespalten wird.

Dabei wird zunächst ein Proteinextrakt des gewünschten Gewebe hergestellt, dem dann das Substrat der GUS zugesetzt wird. Das Substrat ist erst nach der Umsetzung durch GUS fluorimetrisch messbar. Zu verschiedenen Zeitpunkten werden Proben entnommen, die anschließend im Fluorimeter gemessen werden. Dieser Test

wurde mit Leinembryonen verschiedener Altersstadien durchgeführt (21, 24 oder 30 Tage nach Beginn der Blüte, daf = days after flowering). Dazu wurde je ein Embryo in einem 2 ml-Reaktionsgefäß mit Hilfe einer Schwingmühle (Retsch MM 2000) in flüssigem Stickstoff zu Pulver zerrieben. Nach Zugabe von 100 ml EGL-Puffer wurde für 10 min bei 25°C und 14000 x g zentrifugiert. Der Überstand wurde abgenommen und ein zweites Mal zentrifugiert. Wieder wurde der Überstand in ein neues Reaktionsgefäß überführt und bis zur weiteren Verwendung auf Eis gehalten. Von diesem Proteinextrakt wurden 25 ml mit 65 ml EGL-Puffer (ohne DTT) versetzt und für den GUS-Assay eingesetzt. Nun wurden 10 ml des Substrates MUG (10 mM 4-Methyl-umbelliferyl-β-D-glucuronid) dazugegeben, vortex und sofort 30 ml als Nullwert entnommen und mit 470 ml Stopp-Puffer (0,2 M Na<sub>2</sub>CO<sub>3</sub>) versetzt. Dieser Vorgang wurde für alle Proben in einem Abstand von 30 s wiederholt. Die entnommenen Proben wurden bis zur Messung im Kühlschrank gelagert. Weitere Messwerte wurden nach 1 h und nach 2 h entnommen. Für die Messung im Fluorimeter wurde eine Eichreihe erstellt, die Konzentrationen von 0,1 mM bis 10 mM MU (4-Methyl-umbelliferon) enthielt. Waren die Probenwerte außerhalb dieser Konzentrationen, wurde weniger Proteinextrakt eingesetzt (10 ml, 1 ml, 1 ml aus 1:10 Verdünnung), und es wurden kürzere Zeitabstände gemessen (0 h, 30 min, 1 h). Die Messung erfolgte bei einer Excitation von 365 nm und einer Emission von 445 nm in einem Fluoroscan II-Gerät (Labsystem). Alternativ kann die Substratspaltung unter alkalischen Bedingungen fluorometrisch verfolgt werden (Anregung bei 365 nm, Messung der Emission bei 455 nm; Spectrofluorimeter BMG Polarstar+) wie beschrieben in Bustos M.M. et al., 1989 Plant Cell 1:839-853. Alle Proben wurden einer Proteinkonzentrationsbestimmung nach Bradford (1976) unterzogen, um so eine Aussage über die Promotoraktivität und -stärke in verschiedenen Geweben und Pflanzen erlauben.

## EGL-Puffer

35 0,1 M	KPO <sub>4</sub> , pH 7,8
1 mM	EDTA
5 %	Glycerin
1 M	DTT

40 Als weitere Beispiele für Reportergene, die äquivalent benutzt werden können, seien beispielhaft das grün fluoreszierende Protein (GFP) und dessen Derivate genannt (C.Reichel et al. (1996) Proc.Natl.Acad.Sci.USA 93, 5888-5893 und J.Sheen et al., (1995) Plant Journal 8, 777-784) und verschiedene Luciferasen (A.Millar et al. (1992) Plant Mol. Biol. Reporter 10, 324-414). Die ent-

sprechenden Detektionsmethoden sind dem Fachmann bekannt und z.B. genannter Literatur beschrieben.

Beispiele für Promoter-Reportergen-Konstrukte für oben genannte  
5 Promotoren sind im folgenden gegeben. Von diesen Promotoren können Fragmente mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert werden.

10 Folgende Oligonukleotide können beispielsweise verwendet werden:

LeB4 vorne: GAAAGCTTCTCGAGTTATGCATTTCTT

LeB4 hinten: GGGTCTAGATCTGTGACTGTGATAG

DC3a vorne: AGTGGATCCCCGAGCTAACCACAAC

15 DC3a hinten: ATAAGCTTTTCTTTGCAGA

napinvorne: GAAAGCTTCTAATATGATAAACTCTG

napinhinten: GGGTCTAGAAACACATACAAACATCAC

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind  
20 allgemein literaturbekannt.

In einem ersten Schritt werden die Promotorfragmente über PCR amplifiziert, mit geeigneten Restriktionsenzymen geschnitten und in die obigen Kassetten einkloniert. Beispielsweise wird das  
25 LeB4(700)-PCR-Fragment mit XbaI und HindIII geschnitten und in den Vektor pGPTV in die HindIII und XbaI-Schnittstellen 5' vor dem GUS Reportergen kloniert. Das PCR-amplifizierte DC3-Promoterfragment kann mit BamHI und HindIII geschnitten, z.B. in pBluescript (Stratagene) subkloniert und dann in pGPTV in geeignete  
30 Schnittstellen vor das GUS-Reportergen kloniert werden.

Beispielsweise kann ein mit obigen Primern PCR-amplifiziertes napin-Promotorfragment mit einer Größe von 1055bp nach Verdau mit HindIII und XbaI in pGPTV vor das GUS Reportergen einkloniert  
35 werden. Ein äquivalentes Konstrukt könnte ein Napin-Promoterfragment von 1100bp 5'-kloniert vor ein GUS Reportergen mit Intron mit in 3' Richtung nachfolgendem Nos-Terminator in einem pHL9000-Vektor (Hausmann & Töpfer, 1999) sein.

40 Unter Verwendung von oben beschriebenen oder äquivalenten Konstrukten nach Transformation in Lein der Sorte Flanders, läßt sich die GUS-Aktivität in transgenen Leinembryonen verschiedener Altersstadien messen, die mit einem der folgenden Konstrukte transformiert wurden: Napin-GUS, 35S-GUS, LeB4-GUS, USP-GUS. Die  
45 Werte sind Mittelwerte aus ein bis fünf Messungen mit verschiede-

## 91

nen Proteinmengen. Pro Konstrukt wurden je drei Embryonen quantitativ analysiert.

In der quantitativen Analyse ergaben sich große Unterschiede zwischen den verschiedenen samenspezifischen Promotoren. Der Napin Promotor aus *Brassica napus* erwies sich als um zwei bis drei Zehnerpotenzen weniger aktiv als die beiden Promotoren aus *Vicia faba* (LeB4 und USP). Die GUS-Aktivität nahm in der Reihenfolge Napin, LeB4 und USP zu. Die Positivkontrolle 35S bewegte sich in ihrer Aktivität zwischen LeB4 und USP, wohingegen die Negativkontrolle, der nicht transformierte Wildtyp (Sorte Flanders), so gut wie keine Aktivität besaß. Die folgende Tabelle gibt die Mittelwerte der Aktivitäten in den einzelnen Altersstadien sowie gesamt für jedes Konstrukt wieder.

15

Tabelle 3.4: Übersicht über die mittleren GUS-Aktivitäten von Leinembryonen, transformiert mit verschiedenen GUS-Konstrukten. daf: Tage nach Beginn der Blüte.

20 Mit \* markiert wurden Werte, die nur einer Messung zugrundeliegen.

25	GUS-Konstrukt	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	GUS-Aktivität [nmol/h/mg Protein]	% vs. 35S
		Mittelwert 21 daf	Mittelwert 24 daf	Mittelwert 30 daf	Gesamtmittelwert	
30	Ohne	0,06	0,06	*0,02	0,05	0,0007
	35S	649,00	913,00	639,00	734,00	100,00
	Napin	7,70	5,70	3,70	5,70	0,80
	LeB4	1778,00	253,00	283,00	771,00	105,00
	USP	2843,00	3770,00	2107,00	2907,00	396,00

35

### Beispiel 13: Plasmide für die Pflanzentransformation

Zur Pflanzentransformation können binäre Vektoren, wie pBinAR (Höfgen und Willmitzer, Plant Science 66 (1990) 221-230) oder pGPTV (Becker et al 1992, Plant Mol. Biol. 20:1195-1197) oder Derivate davon verwendet werden. Die Konstruktion der binären Vektoren kann durch Ligation der cDNA in Sense- oder Antisense-Orientierung in T-DNA erfolgen. 5' der cDNA aktiviert ein Pflanzenpromotor die Transkription der cDNA. Eine Polyadenylierungssequenz befindet sich 3' von der cDNA. Die binären Vektoren können unterschiedliche Markergene tragen. Insbesondere kann das nptII-Markergen codierend für Kanamycin-Resistenz vermittelt

- durch Neomycinphosphotransferase gegen die herbizidresistente Form eines Acetolactat Synthasegens (Abkürzung: AHAS oder ALS) ausgetauscht werden. Das ALS-Gen ist beschrieben in Ott et al., J. Mol. Biol. 1996, 263:359-360. Der v-ATPase-c1-Promotor kann in
- 5 das Plasmid pBin19 oder pGPTV kloniert werden und durch Klonierung vor das ALS Codierregion für die Markergenexpression genutzt werden. Der genannte Promotor entspricht einem 1153 Basenpaarfragment aus beta-Vulgaris (Plant Mol Biol, 1999, 39:463-475). Dabei können sowohl Sulphonylharnstoffe als auch Imidazolinone
- 10 wie Imazethapyr oder Sulphonylharnstoffe als Antimetaboliten zur Selektion verwendet werden.

- Die gewebespezifische Expression lässt sich unter Verwendung eines gewebespezifischen Promotors erzielen. Beispielsweise kann
- 15 die samenspezifische Expression erreicht werden, indem der DC3- oder der LeB4- oder der USP-Promotor oder der Phaseolin-Promotor 5' der cDNA einkloniert wird. Auch jedes andere samenspezifische Promotorelement wie z.B. der Napin- oder Arcelin Promotor Goossens et al. 1999, Plant Phys. 120(4):1095-1103 und Gerhardt
- 20 et al. 2000, Biochimica et Biophysica Acta 1490(1-2):87-98) kann verwendet werden. Zur konstitutiven Expression in der ganzen Pflanze lässt sich der CaMV-35S-Promotor oder ein v-ATPase C1 Promotor verwenden.

- 25 Insbesondere lassen sich Gene codierend für Desaturasen und Elongasen durch Konstruktion mehrerer Expressionskassetten hintereinander in einen binären Vektor klonieren, um den Stoffwechselweg in Pflanzen nachzubilden.
- 30 Innerhalb einer Expressionskassette kann das zu exprimierende Protein unter Verwendung eines Signalpeptids, beispielsweise für Plastiden, Mitochondrien oder das Endoplasmatische Retikulum, in ein zelluläres Kompartiment dirigiert werden (Kermode, Crit. Rev. Plant Sci. 15, 4 (1996) 285-423). Das Signalpeptid wird 5' im
- 35 Leseraster mit der cDNA einkloniert, um die subzelluläre Lokalisierung des Fusionsprotein zu erreichen.

Beispiele für Multiexpressionskassetten sind im folgenden gegeben.

40

#### I.) Promotor-Terminator-Kassetten

- Expressionskassetten bestehen aus wenigstens zwei funktionellen Einheiten wie einem Promotor und einem Terminator. Zwischen
- 45 Promotor und Terminator können weitere gewünschte Gensequenzen wie Targetting-Sequenzen, Codierregionen von Genen oder Teilen davon etc. eingefügt werden. Zum Aufbau von Expressionskassetten

## 93

werden Promotoren und Terminatoren (USP Promotor: Baeumlein et al., Mol Gen Genet, 1991, 225 (3):459-67); OCS Terminator: Gielen et al. EMBO J. 3 (1984) 835ff.) mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert.

Folgende Oligonukleotide können beispielsweise verwendet werden:

USP1 vorne: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP2 vorne: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

10 USP3 vorne: CCGGAATTCGGCGCGCCGAGCTCCTCGAGCAAATTTACACATTGCCA

USP1 hinten: AAAACTGCAGGCGGCCGCCACCGCGGTGGGCTGGCTATGAAGAAATT

USP2 hinten: CGCGGATCCGCTGGCTATGAAGAAATT

USP3 hinten: TCCCCCGGGATCGATGCCGCGCAGATCTGCTGGCTATGAAGAAATT

OCS1 vorne: AAAACTGCAGTCTAGAAGGCCTCCTGCTTTAATGAGATAT

15 OCS2 vorne: CGCGGATCCGATATCGGGCCCGCTAGCGTTAACCCTGCTTTAATGAGATAT

OCS3 vorne: TCCCCCGGGCCATGGCCTGCTTTAATGAGATAT

OCS1 hinten: CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTCGACGGACAATCAGTAAATTGA

OCS2 hinten: CCCAAGCTTGGCGCGCCGAGCTCGAATTCGTCGACGGACAATCAGTAAATTGA

OCS3 hinten: CCCAAGCTTGGCGCGCCGAGCTCGTCGACGGACAATCAGTAAATTGA

20

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind allgemein literaturbekannt.

In einem ersten Schritt werden ein Promotor und ein Terminator über PCR amplifiziert. Dann wird der Terminator in ein Empfängerplasmid kloniert und in einem zweiten Schritt der Promotor vor den Terminator inseriert. Mithin erhält man eine Expressionskassette auf einem Trägerplasmid. Auf Basis des Plamides pUC19 werden die Plasmide pUT1, pUT2 und pUT3 erstellt.

30

Die Konstrukte sind erfindungsgemäß in SEQ ID NO: 13, 14 und 15 definiert. Sie enthalten auf Basis von pUC19 den USP-Promotor und den OCS Terminator. Auf Basis dieser Plasmide wird das Konstrukt pUT12 erstellt, indem pUT1 mittels SalI/ScaI geschnitten wird und pUT2 mittels XhoI/ScaI geschnitten wird. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung ampicillinresistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die zwei Expressionskassetten enthalten. Die XhoI/SalI Ligation kompatibler Enden hat dabei die beiden Schnittstellen XhoI und SalI zwischen den Expressionskassetten eliminiert. Es resultiert das Plasmid pUT12, das in SEQ ID NO: 16 definiert ist. Anschließend wird pUT12 wiederum mittels Sal/ScaI geschnitten und pUT3 mittels XhoI/ScaI geschnitten. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung ampicillinresistenter Kolonien DNA präpariert

## 94

und per Restriktionsanalyse solche Klone identifiziert, die drei Expressionskassetten enthalten. Auf diese Weise wird ein Set von Multiexpressionskassetten geschaffen, dass für die Insertion gewünschter DNA genutzt werden kann und in Tabelle 3 beschrieben wird und zudem noch weitere Expressionskassetten aufnehmen kann.

Diese enthalten folgende Elemente:

Tabelle 3

10

pUC19-Derivat	Schnittstellen vor dem USP Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
pUT1	EcoRI/AscI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI	SalI/EcoRI/ SacI/AscI/ HindIII
15 pUT2	EcoRI/AscI/ SacI/XhoI	BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AscI/ HindIII
pUT3	EcoRI/AscI/ SacI/XhoI	BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/ AscI/HindIII
pUT12 Doppel-expressionskassette	EcoRI/AscI/ SacI/XhoI	BstXI/NotI/ PstI/XbaI/StuI Und BamHI/EcoRV/ ApaI/NheI/ HpaI	SalI/EcoRI/ SacI/AscI/ HindIII
20 pUT123 Tripel-expressionskassette	EcoRI/AscI/ SacI/XhoI	1.BstXI/NotI/ PstI/XbaI/StuI und 2.BamHI/EcoRV/ ApaI/NheI/ HpaI und 3.BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII

25

Weiterhin lassen sich wie beschrieben und wie in Tabelle 4 näher spezifiziert weitere Multiexpressionskassetten mithilfe des

- 30 i) USP-Promotors oder mithilfe des  
 ii) ca. 700 Basenpaare 3'-Fragmentes des LeB4-Promotors oder mithilfe des  
 iii) DC3-Promotors erzeugen und für samenspezifische Genexpression einsetzen.

35

Der DC3-Promotor ist beschrieben bei Thomas, Plant Cell 1996, 263:359-368 und besteht lediglich aus der Region -117 bis +27 weshalb er mithin einer der kleinsten bekannten samenspezifischen Promotoren darstellt.

40

Von diesen Promotoren können Fragmente mithilfe der Polymerasekettenreaktion isoliert und mit flankierenden Sequenzen nach Wahl auf Basis von synthetischen Oligonukleotiden maßgeschneidert werden.

45

Folgende Oligonukleotide können beispielsweise verwendet werden:

## 95

LeB4 vorne: GAAAGCTTCTCGAGTTATGCATTTCTT  
LeB4 hinten: GGGTCTAGATCTGTGACTGTGATAG  
DC3a vorne: CCGGAATTCGGCGCGCCGAGCTCCTCGAG  
DC3a hinten: CGCGGATCCTAGCTTTTCTTGGCAGATG

5

Die Methoden sind dem Fachmann auf dem Gebiet bekannt und sind allgemein literaturbekannt.

In einem ersten Schritt werden die Promotorfragmente über PCR amplifiziert, mit geeigneten Restriktionsenzymen geschnitten und in die obigen Kassetten einkloniert. Beispielsweise wird das LeB4(700)-PCR-Fragment mit XhoI und BglII geschnitten und in die XhoI und BglII-Schnittstellen des Plasmids pUT3 eingesetzt, um pLT3 zu erhalten.

15

Vorteilhafte Expressionskassetten enthalten auf Basis von pUC19 (Vieira und Messing (1982); Gene 19, 259), die SEQ ID NO: 32, den LeB4-Promotor und die Sequenzen SEQ ID NO: 33, SEQ ID NO: 34 oder SEQ ID NO: 35. Auf Basis dieser Plasmide wird das Konstrukt pLT12  
20 erstellt, indem pUT1 mittels SalI/ScaI geschnitten wird und pUT2 mittels XhoI/ScaI geschnitten wird. Die die Expressionskassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung Ampicillin-resistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone  
25 identifiziert, die zwei Expressionskassetten enthalten. Die XhoI/SalI Ligation kompatibler Enden hat dabei die beiden Schnittstellen XhoI und SalI zwischen den Expressionskassetten eliminiert. Es resultiert das Plasmid pUT12, das in SEQ ID NO: 16 definiert ist. Anschließend wird pUT12 wiederum mittels Sal/ScaI geschnitten und pUT3 mittels XhoI/ScaI geschnitten. Die die Expressions-  
30 kassetten enthaltenden Fragmente werden ligiert und in E. coli XLI blue MRF transformiert. Es wird nach Vereinzelung Ampicillin-resistenter Kolonien DNA präpariert und per Restriktionsanalyse solche Klone identifiziert, die drei Expressionskassetten enthal-  
35 ten. Auf diese Weise wird eine Auswahl von Multiexpressionskassetten geschaffen, die für die Insertion gewünschter DNA genutzt werden kann und in Tabelle 3 beschrieben wird und zudem noch weitere Expressionskassetten aufnehmen kann.

40 Die Expressionskassetten können mehrfach den selben Promotor enthalten oder aber über drei verschiedene Promotoren aufgebaut werden.

45

Tabelle 4: Multiple Expressionskassetten

	Plasmidname des pUC19-Derivates	Schnittstellen vor dem jeweiligen Promotor	Multiple Klonierungs-Schnittstellen	Schnittstellen hinter dem OCS-Terminator
5	pUT1 (pUC19 mit USP-OCS1)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/PstI/ XbaI/StuI	SalI/EcoRI/SacI/AscI/HindIII
	pDCT (pUC19 mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(2) BamHI/EcoRV/ ApaI/NheI/HpaI	SalI/EcoRI/SacI/AscI/HindIII
10	pLeBT (pUC19-mit LeB4(700)-OCS)	EcoRI/AscI/SacI/XhoI	(3) BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII
15	pUD12 (pUC 19 mit mit USP-OCS1 und mit DC3-OCS)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/EcoRV/ ApaI/NheI/HpaI	SalI/EcoRI/SacI/AscI/HindIII
20	pUDL123 Triple expression cassette (pUC19 mit USP/ DC3 und LeB4-700)	EcoRI/AscI/SacI/XhoI	(1) BstXI/NotI/ PstI/XbaI/StuI und (2) BamHI/ (EcoRV*)/ApaI/NheI/HpaI und (3) BglII/NaeI/ ClaI/SmaI/NcoI	SalI/SacI/AscI/HindIII

\* EcoRV Schnittstelle schneidet im 700 Basenpaarfragment des LeB4 Promotors (LeB4-700)

25

Analog lassen sich weitere Promotoren für Multigenkonstrukte erzeugen insbesondere unter Verwendung des

30 a) 2,4 kB Fragmentes des LeB4-Promotors (Bäumlein et al., 1991: Mol.Gen.Genet. 225,121-128) oder mithilfe des

b) Phaseolin-Promotors (Bustos et al. (1989) Plant Cell 1,839-853) oder mithilfe des

35 c) konstitutiven v-ATPase c1-Promotors.

Es kann insbesondere wünschenswert sein, weitere besonders geeignete Promotoren zum Aufbau samenspezifischer Multiexpressionskassetten wie z.B. eines der Fragmente des Napin-Promotors (Stal-  
40 berg et al., 1993: Plant Mol.Biol.23,671-683) oder den Arcelin-5 Promotor (A.Goossens et al., 1999: plant Physiol. 120,1095-1104) zu verwenden.

45

## 97

- ii) Erstellung von Expressionskonstrukten in pUC19- oder pGPTV Derivaten, die Promotor und Terminator erhalten und in Kombination mit gewünschten Gensequenzen zur PUFA Gen-expression in pflanzlichen Expressionskassetten enthalten.

5

- Multiexpressionskassetten können mittels AscI direkt von pUC19-Derivaten aus Tabelle 3 in den Vektor pGPTV+AscI (siehe iii.)) über die AscI Schnittstelle inseriert werden und stehen zur Inserierung von Zielgenen zur Verfügung. Die entsprechenden
- 10 Genkonstrukte (pBUT1 ist in SEQUENZ ID NO: 20, pBUT2 ist in SEQUENZ ID NO: 21, pBUT 3 ist in SEQUENZ ID NO: 22, pBUT12 ist in SEQUENZ ID NO: 22 und pBUT123 ist in SEQUENZ ID NO: 24 dargestellt) stehen erfindungsgemäß als Kit zur Verfügung. Alternativ können Gensequenzen in die pUC19 basierten
- 15 Expressionskassetten inseriert werden und als AscI Fragment in pGPTV+AscI eingesetzt werden.

- In pUT12 wird zunächst über BstXI und XbaI die D-6-Elongase Pp\_PSE1 in die erste Kassette inseriert. Dann wird die
- 20 D-6-Desaturase aus Moos (Pp\_des6) über BamHI/NaeI in die zweite Kassette inseriert. Es entsteht das Konstrukt pUT-ED. Das AscI Fragment aus dem Plasmid pUT-ED wird in den mit AscI geschnittenen Vektor pGPTV+AscI inseriert und die Orientierung des inserierten Fragmentes mittels Restriktion oder Sequenzierung
- 25 ermittelt. Es entsteht das Plasmid pB-DHGLA, dessen vollständige Sequenz in SEQUENZ ID NO. 25 dargestellt ist. Die Codierregion der Physcomitrella delta 6 Elongase ist in SEQUENZ ID NO. 26 dargestellt, die der delta 6 Desaturase aus Physcomitrella in SEQUENZ ID NO: 27.

30

- In pUT123 wird zunächst über BstXI und XbaI die  $\Delta$ -6-Elongase Pp\_PSE1 in die erste Kassette inseriert. Dann wird die  $\Delta$ -6-Desaturase aus Moos (Pp\_des6) über BamHI/NaeI in die zweite Kassette inseriert und schließlich die  $\Delta$ -5-Desaturase
- 35 aus Phaeodactylum (Pt\_des5) über BglII in die dritte Kassette inseriert. Das Dreifachkonstrukt erhält den Namen pARA1. Unter Berücksichtigung sequenzspezifischer Restriktionsschnittstellen können weitere Expressionskassetten gemäß Tabelle 5 mit der Bezeichnung pARA2, pARA3 und pARA4 erstellt werden.

40

- Das AscI Fragment aus dem Plasmid pARA1 wird in den mit AscI geschnittenen Vektor pGPTV+AscI inseriert und die Orientierung des inserierten Fragmentes mittels Restriktion oder Sequenzierung ermittelt. Die vollständige Sequenz des resultierenden Plasmides
- 45 pBARA1 ist in SEQUENZ ID NO. 28 dargestellt. Die Codierregion der Physcomitrella delta 6 Elongase ist in SEQUENZ ID NO. 29 dargestellt, die der delta 6 Desaturase aus Physcomitrella

in SEQUENZ ID NO: 30 und die der delta-5 Desaturase aus *Phaeodactylum tricornutum* in SEQUENZ ID NO: 31.

Tabelle 5: Kombinationen von Desaturasen und Elongasen

5

	Gen Plasmid	$\Delta$ -6-Desaturase	$\Delta$ -5-Desaturase	$\Delta$ -6-Elongase
1	PUT-ED	Pp_des6	---	Pp_PSE1
2	pARA1	Pt_des6	Pt_des5	Pp_PSE1
3	pARA2	Pt_des6	Ce_des5	Pp_PSE1
10 4	pARA3	Pt_des6	Ce_des5	Pp_PSE1
5	pARA4	Ce_des6	Ce_des5	Ce_PSE1
6	PBDHGLA	Pt_des6	---	Pp_PSE1
7	PBARAI	Pt_des6	Pt_des5	Pp_PSE1

15 Plasmide 1 bis 5 sind pUC Derivate, Plasmide 6 bis 7 sind binäre Pflanzentransformationsvektoren

Pp = *Physcomitrella patens*, Pt = *Phaeodactylum tricornutum*

Pp\_PSE1 entspricht der Sequenz aus SEQ ID NO: 9.

20 PSE = PUFA spezifische  $\Delta$ -6-Elongase

Ce\_des5 =  $\Delta$ -5-Desaturase aus *Caenorhabditis elegans* (Genbank Acc. Nr. AF078796)

Ce\_des6 =  $\Delta$ -6-Desaturase aus *Caenorhabditis elegans elegans* (Genbank Acc. Nr. AF031477, Basen 11-1342)

25 Ce\_PSE1 =  $\Delta$ -6-Elongase aus *Caenorhabditis elegans* (Genbank Acc. Nr. AF244356, Basen 1-867)

Auch weitere Desaturasen oder Elongasegensequenzen können in Expressionskassetten beschriebener Art inseriert werden wie  
30 z.B. Genbank Acc. Nr. AF231981, NM\_013402, AF206662, AF268031, AF226273, AF110510 oder AF110509.

iii) Transfer von Expressionskassetten in Vektoren zur  
Transformation von *Agrobacterium tumefaciens* und zur  
35 Transformation von Pflanzen

Chimäre Genkonstrukte auf Basis der in pUC19 beschriebenen können mittels AscI in den binären Vektor pGPTV inseriert. Die multiple  
Klonierungssequenz wird zu diesem Zweck um eine AscI Schnitt-  
40 stelle erweitert. Zu diesem Zweck wird der Polylinker als zwei doppelsträngige Oligonukleotide neu synthetisiert, wobei eine zusätzliche AscI DNA Sequenz eingefügt wird. Das Oligonukleotid wird mittels EcoRI und HindIII in den Vektor pGPTV inseriert. Es entsteht das Plasmid pGPTV+AscI. Die notwendigen Kloniertechniken  
45 sind dem Fachmann bekannt und können einfach wie in Beispiel 1 beschrieben nachgelesen werden.

Beispiel 14: *In vivo*-Mutagenese

Die *in vivo*-Mutagenese von Mikroorganismen kann mittels Passage der Plasmid- (oder einer anderen Vektor-) DNA durch *E. coli* oder andere Mikroorganismen (z.B. *Bacillus* spp. oder Hefen, wie *Saccharomyces cerevisiae*), bei denen die Fähigkeiten, die Unversehrtheit ihrer genetischen Information aufrechtzuerhalten, gestört ist, erfolgen. Übliche Mutator-Stämme haben Mutationen in den Genen für das DNA-Reparatursystem (z.B. *mthLS*, *mutD*, *mutT* usw.; als Literaturstelle siehe Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, S. 2277-2294, ASM: Washington). Diese Stämme sind dem Fachmann bekannt. Die Verwendung dieser Stämme ist beispielsweise in Greener, A., und Callahan, M. (1994) *Strategies* 7:32-34, erläutert. Der Transfer mutierter DNA-Moleküle in Pflanzen erfolgt vorzugsweise nach Selektion und Test der Mikroorganismen. Transgene Pflanzen werden nach verschiedenen Beispielen im Beispielteil dieses Dokumentes erzeugt.

## Beispiel 15: Untersuchung der Expression eines rekombinanten Genproduktes in einem transformierten Organismus

Die Aktivität eines rekombinanten Genproduktes im transformierten Wirtsorganismus kann auf der Transkriptions- und/oder der Translationsebene gemessen werden.

Ein geeignetes Verfahren zur Bestimmung der Menge an Transkription des Gens (ein Hinweis auf die Menge an RNA, die für die Translation des Genproduktes zur Verfügung steht) ist die Durchführung eines Northern-Blots wie unten ausgeführt (als Bezugsstelle siehe Ausubel et al. (1988) *Current Protocols in Molecular Biology*, Wiley: New York, oder den oben erwähnten Beispielteil), wobei ein Primer, der so gestaltet ist, dass er an das Gen von Interesse bindet, mit einer nachweisbaren Markierung (gewöhnlich radioaktiv oder chemilumineszent) markiert wird, so dass, wenn die Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix transferiert und mit dieser Sonde inkubiert wird, die Bindung und das Ausmaß der Bindung der Sonde das Vorliegen und auch die Menge der mRNA für dieses Gen anzeigt. Diese Information zeigt den Grad der Transkription des transformierten Gens an. Zelluläre Gesamt-RNA kann aus Zellen, Geweben oder Organen mit mehreren Verfahren, die alle im Fachgebiet bekannt sind, wie zum Beispiel das von Bormann, E.R., et al. (1992) *Mol. Microbiol.* 6:317-326 beschriebene, präpariert werden.

## Northern-Hybridisierung

Für die RNA-Hybridisierung wurden 20 µg Gesamt-RNA oder 1 µg poly(A)<sup>+</sup>-RNA mittels Gelelektrophorese in Agarosegelen mit  
5 einer Stärke von 1,25 % unter Verwendung von Formaldehyd, wie beschrieben in Amasino (1986, Anal. Biochem. 152, 304) aufgetrennt, mittels Kapillaranziehung unter Verwendung von 10 x SSC auf positiv geladene Nylonmembranen (Hybond N+, Amersham, Braunschweig) übertragen, mittels UV-Licht immobilisiert und  
10 3 Stunden bei 68°C unter Verwendung von Hybridisierungspuffer (10 % Dextransulfat Gew./Vol., 1 M NaCl, 1 % SDS, 100 mg Heringsperma-DNA) vorhybridisiert. Die Markierung der DNA-Sonde mit dem Highprime DNA labeling-Kit (Roche, Mannheim, Deutschland) erfolgte während der Vorhybridisierung unter Verwendung von  
15 alpha-<sup>32</sup>P-dCTP (Amersham, Braunschweig, Deutschland). Die Hybridisierung wurde nach Zugabe der markierten DNA-Sonde im gleichen Puffer bei 68°C über Nacht durchgeführt. Die Waschschriffe wurden zweimal für 15 min unter Verwendung von 2 X SSC und zweimal für 30 min unter Verwendung von 1 X SSC, 1 % SDS, bei 68°C durch-  
20 geführt. Die Exposition der verschlossenen Filter wurde bei -70°C für einen Zeitraum von 1 bis 14 T durchgeführt.

Zur Untersuchung des Vorliegens oder der relativen Menge an von dieser mRNA translatiertem Protein können Standardtechniken, wie  
25 ein Western-Blot, eingesetzt werden (siehe beispielsweise Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). Bei diesem Verfahren werden die zellulären Gesamtproteine extrahiert, mittels Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrozellulose, übertragen und mit einer  
30 Sonde, wie einem Antikörper, der spezifisch an das gewünschte Protein bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszenten oder kolorimetrischen Markierung versehen, die sich leicht nachweisen lässt. Das Vorliegen und die Menge der beobachteten Markierung zeigt das Vorliegen und die Menge des  
35 gewünschten, in der Zelle vorliegenden mutierten Proteins an.

Beispiel 16: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

40 Die Auswirkung der genetischen Modifikation in Pflanzen, Pilzen, Algen, Ciliaten oder auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierten Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen)  
45 gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die erhöhte Produktion des gewünschten Produktes (d.h. von Lipiden oder einer Fettsäure) untersucht wird. Diese Analyse-

## 101

- techniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie
- 5 (siehe beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehn et al. (1993) Biotechnology, Bd. 3, Kapitel III:
- 10 "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988)
- 15 Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).
- 20 Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145, beschrieben extrahiert. Die qualitative und quantitative Lipid- oder Fettsäureanalyse ist beschrieben
- 25 bei Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford:
- 30 Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

- Zusätzlich zur Messung des Endproduktes der Fermentation ist es auch möglich, andere Komponenten der Stoffwechselwege zu ana-
- 35 lysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere
- 40 Ionen), Messungen der Biomassezusammensetzung und des Wachstums, Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F.
- 45 Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192

## 102

(ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME, 5 Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Massenspektrometrie; TAG, Triacylglycerin; TLC, Dünnschichtchromatographie).

Der unzweideutige Nachweis für das Vorliegen von Fettsäure-  
10 produkten kann mittels Analyse rekombinanter Organismen nach Standard-Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierte Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, Gaschromato-  
15 graphie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das  
20 Material muss nach dem Aufbrechen zentrifugiert werden. Das Sediment wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl-  
25 und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester werden in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 Mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min  
30 und 5 min bei 240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester muss unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definiert werden.

35 Bei Fettsäuren, für die keine Standards verfügbar sind, muss die Identität über Derivatisierung und anschließende GC-MS-Analyse gezeigt werden. Beispielsweise muss die Lokalisierung von Fettsäuren mit Dreifachbindung über GC-MS nach Derivatisierung mit 4,4-Dimethoxyoxazolin-Derivaten (Christie, 1998, siehe oben)  
40 gezeigt werden.

## 103

## Expressionskonstrukte in heterologen mikrobiellen Systemen

## Stämme, Wachstumsbedingungen und Plasmide

- 5 Der *Escherichia coli*-Stamm XL1 Blue MRF' kan (Stratagene) wurde zur Subklonierung der neuen Desaturase pPDesaturasel aus *Physcomitrella patens* verwendet. Für die funktionelle Expression dieses Gens verwendeten wir den *Saccharomyces cerevisiae*-Stamm INVSc 1 (Invitrogen Co.). *E. coli* wurde in Luria-Bertini-Brühe (LB,
- 10 Duchefa, Haarlem, Niederlande) bei 37°C kultiviert. Wenn nötig, wurde Ampicillin (100 mg/Liter) zugegeben, und 1,5 % Agar (Gew./Vol.) wurde für feste LB-Medien hinzugefügt. *S. cerevisiae* wurde bei 30°C entweder in YPG-Medium oder in komplettem Minimalmedium ohne Uracil (CMDm; siehe in: Ausubel, F.M., Brent, R., Kingston,
- 15 R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L.B., Coen, D.M., und Varki, A. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York) mit entweder 2 % (Gew./Vol.) Raffinose oder Glucose kultiviert. Für feste Medien wurden 2 % (Gew./Vol.) Bacto™-Agar (Difco)
- 20 hinzugefügt. Die zur Klonierung und Expression verwendeten Plasmide sind pUC18 (Pharmacia) und pYES2 (Invitrogen Co.).

Beispiel 17: Klonierung und Expression PUFA-spezifischer  
Desaturasen aus *Phaeodactylum tricornutum*

25

- Für die Expression in Hefe wurden die *Phaeodactylum tricornutum* -cDNA-Klone aus Seq ID NO: 1, 3, 5 oder 11 bzw. die Sequenzen aus SEQ ID NO: 7 oder 9 bzw andere gewünschte Sequenzen zuerst so modifiziert, dass lediglich die Codierregion mittels Polymerase
- 30 Kettenreaktion unter Zuhilfenahme zweier Oligonukleotide amplifiziert werden. Dabei wurde darauf geachtet, dass eine Konsensussequenz vor dem Startcodon zur effizienten Translation eingehalten wurde. Entweder wurde hierzu die Basenfolge ATA oder AAA gewählt und vor das ATG in die Sequenz eingefügt (Kozak, M.
- 35 (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, Cell 44, 283-292). Vor diesem Konsensustriplett wurde zusätzlich eine Restriktionsschnittstelle eingeführt, die kompatibel sein muss zur Schnittstelle des Zielvektors,
- 40 in den das Fragment kloniert werden soll und mit dessen Hilfe die Genexpression in Mikroorganismen oder Pflanzen erfolgen soll.

## 104

Die PCR-Reaktion wurde mit Plasmid-DNA als Matrize in einem Thermocycler (Biometra) mit der Pfu-DNA-(Stratagene) Polymerase und dem folgenden Temperaturprogramm durchgeführt: 3 min bei 96°C, gefolgt von 30 Zyklen mit 30 s bei 96°C, 30 s bei 55°C und 2 min bei 72°C, 1 Zyklus mit 10 min bei 72°C und Stop bei 4°C. Die Anlagerungstemperatur wurde je nach gewählten Oligonukleotiden variiert. Pro Kilobasenpaare DNA ist von einer Synthesezeit von etwa einer Minute auszugehen. Weitere Parameter, die Einfluss auf die PCR haben wie z.B. Mg-Ionen, Salz, DNA Polymerase etc., sind dem Fachmann auf dem Gebiet geläufig und können nach Bedarf variiert werden.

Die korrekte Größe des amplifizierten DNA-Fragments wurde mittels Agarose-TBE-Gelelektrophorese bestätigt. Die amplifizierte DNA wurde aus dem Gel mit dem QIAquick-Gelextraktionskit (QIAGEN) extrahiert und in die SmaI-Restriktionsstelle des dephosphorylierten Vektors pUC18 unter Verwendung des Sure Clone Ligations Kit (Pharmacia) ligiert, wobei die pUC-Derivate erhalten wurden. Nach der Transformation von E. coli XL1 Blue MRF' kan wurde eine DNA-Minipräparation (Riggs, M.G., & McLachlan, A. (1986) A simplified screening procedure for large numbers of plasmid mini-preparation. BioTechniques 4, 310-313) an ampicillinresistenten Transformanten durchgeführt, und positive Klone mittels BamHI-Restriktionsanalyse identifiziert. Die Sequenz des klonierten PCR-Produktes wurde mittels Resequenzierung unter Verwendung des ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt) bestätigt.

Δ5 Acyl Lipid desaturase, Pt\_des5

Primer 1 GAG CTC ACA TAA TGG CTC CGG ATG CGG ATA AGC  
Primer 2 CTC GAG TTA CGC CCG TCC GGT CAA GGG

Das PCR-Fragment (1428bp) wurde mithilfe des Sure Clone Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment SacI/XhoI verdaut und das Fragment mithilfe entsprechender Restriktionsschnittstellen in pYES2 oder pYES6 inseriert.

Δ6 Acyl Lipid desaturase, Pt\_des6

Primer 3 GGA TCC ACA TAA TGG GCA AAG GAG GGG ACG CTC GGG  
Primer 4 CTC GAG TTA CAT GGC GGG TCC ATC GGG

Das PCR-Fragment (1451 bp) wurde mithilfe des Sure Clone Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment BamHI/XhoI verdaut und das Fragment mithilfe entsprechender Restriktionsschnittstellen in pYES2 oder pYES6 inseriert.

## 105

$\Delta$ 12 Acyl Lipid desaturase, Pt\_des12

Primer 5 GGA TCC ACA TAA TGG TTC GCT TTT CAA CAG CC

Primer 6 CTC GAG TTA TTC GCT CGA TAA TTT GC

5  $\Delta$ 12 Acyl Lipid desaturase, Pt\_des12.2

Primer 7 GGA TCC ACA TAA TGG GTA AGG GAG GTC AAC G

Primer 8 CTC GAG TCA TGC GGC TTT GTT TCG C

- Das PCR Fragment (1505bp) wurde mithilfe des Sure Clone  
10 Kit (Pharmacia) in pUC 18 kloniert, das inserierte Fragment  
BamHI/XhoI verdaut und das Fragment mithilfe entsprechender  
Restriktionsschnittstellen in pYES2 oder pYES6 inseriert.

- Die Plasmid-DNA wurde mit Restriktionsenzym/en passend zur  
15 eingeführten Schnittstelle der Primersequenz gespalten und  
das erhaltene Fragment in die kompatiblen Restriktionsstellen  
des dephosphorylierten Hefe-E. coli-Shuttlevektors pYES2 oder  
pYES6 ligiert, wobei pYES-Derivate erhalten werden. Nach der  
Transformation von E. coli und DNA-Minipräparation aus den  
20 Transformanden wurde die Orientierung des DNA-Fragments im  
Vektor durch geeignete Restriktionsspaltung oder Sequenzierung  
überprüft. Ein Klon wurde für die DNA-Maxipräparation mit dem  
Nucleobond® AX 500 Plasmid-DNA-Extraktionskit (Macherey-Nagel,  
Düringen) angezogen.

- 25 *Saccharomyces cerevisiae* INVSc1 wurde mit den pYES-Derivaten  
und pYES Leervektor mittels eines PEG/Lithiumacetat-Protokolls  
transformiert (Ausubel et al., 1995). Nach der Selektion  
auf CMDum-Agarplatten mit 2 % Glucose wurden pYES-Derivate-  
30 Transformanden und eine pYES2-Transformande zur weiteren Anzucht  
und funktionellen Expression ausgewählt. Bei pYES6-Derivaten  
wurde Blastidicin als Antimetabolit verwendet. Im Fall von  
Coexpressionen auf Basis von pYES2 und pYES6 wurde auf Minimal-  
medium mit Blastidicin selektiert.

- 35 Funktionelle Expression einer Desaturaseaktivität in Hefe

Vorkultur

- 40 20 ml CMDum-Flüssigmedium ohne Uracil aber mit 2 % (Gew./Vol.)  
Raffinose wurden mit den transgenen Hefeklonen (pYES2) angeimpft  
und 3 Tage bei 30°C, 200 rpm gezüchtet, bis eine optische Dichte  
bei 600 nm (OD<sub>600</sub>) von 1,5 bis 2 erreicht wurde. Wurde als Vektor  
pYES6 verwendet, so wurde zusätzlich auf Blastidicin als Anti-  
45 metabolit selektioniert.

## Hauptkultur

Für die Expression wurden 20 ml CMdum-Flüssigmedium ohne Uracil aber mit 2 % Raffinose und 1 % (Vol./Vol.) Tergitol NP-40

- 5 mit Fettsäuresubstraten auf eine Endkonzentration von 0,003 % (Gew./Vol.) angereichert. Die Medien wurden mit den Vorkulturen auf eine OD<sub>600</sub> von 0,05 angeimpft. Die Expression wurde bei einer OD<sub>600</sub> von 0,2 mit 2 % (Gew./Vol.) Galaktose für 16 Std. induziert, wonach die Kulturen eine OD<sub>600</sub> von 0,8-1,2 geerntet wurden.

10

## Fettsäureanalyse

- Die Gesamt-Fettsäuren wurden aus Hefekulturen extrahiert und mittels Gaschromatographie analysiert. Davon wurden Zellen von 5
- 15 ml Kultur mittels Zentrifugation (1000 x g, 10 min, 4°C) geerntet und einmal mit 100 mM NaHCO<sub>3</sub>, pH 8,0, gewaschen, um restliches Medium und Fettsäuren zu entfernen. Zur Herstellung des Fettsäuremethylester (FAMES oder Singular FAME) wurden die Zellsedimente mit 1 M methanolischer H<sub>2</sub>SO<sub>4</sub> und 2 % (Vol./Vol.)
- 20 Dimethoxypropan für 1 Std. bei 80°C behandelt. Die FAMES wurden zweimal mit 2 ml Petrolether extrahiert, einmal mit 100 mM NaHCO<sub>3</sub>, pH 8,0, und einmal mit destilliertem Wasser gewaschen und mit Na<sub>2</sub>SO<sub>4</sub> getrocknet. Das organische Lösungsmittel wurde unter einem Argonstrom verdampft, und die FAMES wurden in 50 mikrol Petrol-
- 25 ether gelöst. Die Proben wurden auf einer ZEBRON-ZB-Wax-Kapillarsäule (30 m, 0,32 mm, 0,25 mikro m; Phenomenex) in einem Hewlett Packard-6850-Gaschromatograph mit einem Flammenionisationsdetektor aufgetrennt. Die Ofentemperatur wurde von 70°C (1 min halten) bis 200°C mit einer Rate von 20°C/min, dann auf 250°C
- 30 (5 min halten) mit einer Rate von 5°C/min und schließlich auf 260°C mit einer Rate von 5°C/min programmiert. Stickstoff wurde als Trägergas verwendet (4,5 ml/min bei 70°C). Die Fettsäuren wurden durch Vergleich mit Retentionszeiten von FAME-Standards (SIGMA) identifiziert.

35

## Expressionsanalyse

- Die Verhältnisse der zugegebenen und aufgenommenen Fettsäuresubstrate wurden ermittelt und so Quantität und Qualität der
- 40 Desaturasereaktion gemäß Tabelle 6, Tabelle 7 und Tabelle 8 erfasst.

Ergebnis der Expression einer *Phaeodactylum tricornutum* Δ-6-Acyl Lipid Desaturase in Hefe:

45

107

Tabelle 6

	Fettsäure	pYes2	pYes2-Ptd6 gefüttert mit		
		-	-	+18:2	+18:3
5	16:0	13,3	18,9	28,4	16,7
	16:1Δ9	45,4	44,7	12,5	16,9
	16:2Δ6,9	-	4,3	-	-
	18:0	4,9	6,3	10,4	9,1
	18:1Δ9	36,4	24,1	6,8	11,8
10	18:2Δ6,9	-	1,8	-	-
	18:2Δ9,12	-	-	33,4	-
	18:3Δ6,12,15	-	-	4,9	-
	18:3Δ9,12,15	-	-	-	43,1
	18:4Δ6,9,12,15	-	-	-	2,3

15 Die Angaben stellen Mol-% entsprechender cis-Fettsäuren dar.

Ergebnis der Expression einer *Phaeodactylum tricornutum* Δ-5-Acyl  
Lipid Desaturase in Hefe:

20 Tabelle 7

	Fett- säure	pYES2		pYES_Ptd5-Konstrukt gefüttert mit						
		Leer	Kon- trolle	18:2	18:3	20:1	20:1	20:2	20:3	20:3
25						Δ8	Δ11	Δ11,14	Ω3	Ω6
	16:0Δ	16,9	20,4	27,7	24,4	16,2	21	17,6	19,5	22,8
	16:1Δ9	44,7	44,1	13,2	9,6	37,4	39,4	38,3	36,9	30,7
	18:0	6,1	6,9	10,54	9,8	4,7	7,9	6,3	6,8	8,2
	18:1Δ9	31,72	28,1	8,77	6	15	26	29,5	25,6	21,1
30	18:2Δ5,9		0,17	0	0	0	0,09	0,21	0,09	9
	18:2Δ9,12	-		39,7	-	-	-	-	-	-
	18:3Δ9,12,15	-			49,9	--	-	-	-	-
	20:1Δ8	-		-	-	25,5	-	-	-	-
	20:1Δ11	-		-	-	-	5,41	-	-	-
35	20:2Δ5,11	-		-	-	-	0,21	-	-	-
	20:2Δ11,14	-		-	-	-	-	6,48	-	-
	20:3Δ5,11,14	-		-	-	-	-	0,76	-	-
	20:3Δ11,14,17	-		-	-	-	-	-	9,83	-
	20:3Δ8,11,14	-		-	-	-	-	-	-	13,69
40	20:4Δ5,11,14,17	-		-	-	-	-	-	1,16	-
	20:4Δ5,8,11,14	-		-	-	-	-	-	-	3,08

Die Angaben stellen Mol-% Fettsäuren von cis-Fettsäuren dar.

## 108

Aus weiteren Fütterungsversuchen wurde gefunden, dass C18:1 $\Delta$ 9 in der Anwesenheit von C18:2 $\Delta$ 9,11 oder C18:3 $\Delta$ 9,12,15 oder C20:1 $\Delta$ 8 Fettsäuren nicht desaturiert wurde während in Anwesenheit von C20:1 $\Delta$ 11, C20:2 $\Delta$ 11,14 und C20:3 $\Delta$ 8,11,14 auch C18:1 desaturiert wird. Ebenfalls keine Desaturierung erfolgte in Anwesenheit von C20:3 $\Delta$ 8,11,14.

Bei Nutzung des Protease-defizienten Hefestammes C13BYS86 (Kunze I. et al., Biochemica et Biophysica Acta (1999) 1410:287-298) für die Expression der  $\Delta$ -5-Desaturase aus *Phaeodactylum tricornutum* auf Vollmedium mit Blastocidin wurde gefunden, dass C20:4  $\Delta$ 8,11,14,17 als Substrat der  $\Delta$ -5-Desaturase mit 20 % Umsatzrate ebenso gut umgesetzt wurde wie C20:3  $\Delta$ 8,11,14. Alternativ können auch die Auxotrophiemarker leu2, ura3 oder his für Genexpression genutzt werden.

In einem weiteren Coexpressionsexperiment von  $\Delta$ -5 Desaturase aus *Phaeodactylum* und  $\Delta$ -6 Elongase aus *Physcomitrella* wurde der Stamm UTL7A (Warnecke et al., J. Biol. Chem. (1999) 274(19):13048-13059) benutzt, wobei die  $\Delta$ -5 Desaturase ca 10 % C20:3  $\Delta$ 8,11,14 zu C20:4  $\Delta$ 5,8,11,14 umsetzte.

Weitere Fütterungsexperimente mit verschiedensten anderen Fettsäuren allein oder in Kombination (z.B. Linolsäure, 20:3  $\Delta$ -5,11,14-Fettsäure, alpha- oder gamma Linolensäure, Stearidonsäure, Arachidonsäure, Eicosapentaensäure etc.) können zur detaillierteren Bestätigung der Substratspezifität und -Selektivität dieser Desaturasen durchgeführt werden.

Tabelle 8: Ergebnis der Coexpression einer *Phaeodactylum tricornutum*  $\Delta$ -5-Acyl Lipid Desaturase und einer  $\Delta$ -6 Elongase aus Moos in Hefe auf Basis der Expressionsvektoren pYes2 und pYes6

	pYes2-Elo		pYes2-Elo and pYes6-Ptd5	
	+18:3	+18:4	+18:3	+18:4
16:0	15,0	14,8	15,6	15,1
16:1 $\Delta$ 9	27,7	29,2	27,5	29,0
18:0	5,6	6,3	5,7	6,4
18:1 $\Delta$ 9	17,1	30,8	27,4	31,6
18:3 $\Delta$ 6,9,12	7,60	-	7,8	-
18:4 $\Delta$ 6,9,12,15	-	6,71	-	6,4
20:3 $\Delta$ 8,11,14	15,92	-	13,55	-
20:4 $\Delta$ 5,8,11,14	-	-	1,31	-
20:4 $\Delta$ 8,11,14,17	-	11,4	-	10,31
20:5 $\Delta$ 5,8,11,14,17	-	-	-	0,53

## 109

Aus den Substratumsetzungen geht hervor, dass die verwendete  $\Delta$ -5-Desaturase aus *Phaeodactylum* und die  $\Delta$ -6-Elongase aus *Physcomitrella patens* bzgl. der Substrataktivität und insbesondere der Substratspezifität geeignet sind, um Arachidon-  
5 säure bzw. Eicosapentaensäure mithilfe erfindungsgemäßer Sequenzen zu produzieren.

Die Fragmentierungsmuster und Massenspektren von DMOX-Derivaten von Standards als auch den Peakfraktionen per GC identifizierter  
10 Fettsäuren der in Tabelle 6, 7 und 8 aufgeführten, zeigen vergleichsweise identische Ergebnisse, wodurch die jeweilige Position der Doppelbindung über die bloße GC-Detektion hinaus abgesichert wurde.

15 Beispiel 18: Reinigung des gewünschten Produktes aus transformierten Organismen

Die Gewinnung des gewünschten Produktes aus Pflanzenmaterial oder Pilzen, Algen, Ciliaten, tierischen Zellen oder aus dem Überstand  
20 der vorstehend beschriebenen Kulturen kann durch verschiedene, im Fachgebiet bekannte Verfahren erfolgen. Wird das gewünschte Produkt nicht aus den Zellen sezerniert, können die Zellen aus der Kultur durch langsame Zentrifugation geerntet werden, die Zellen können durch Standardtechniken, wie mechanische Kraft oder  
25 Ultraschallbehandlung, lysiert werden. Organe von Pflanzen können mechanisch von anderem Gewebe oder anderen Organen getrennt werden. Nach der Homogenisation werden die Zelltrümmer durch Zentrifugation entfernt, und die Überstandsfraktion, welche die löslichen Proteine enthält, wird zur weiteren Reinigung  
30 der gewünschten Verbindung aufbewahrt. Wird das Produkt aus gewünschten Zellen sezerniert, werden die Zellen durch langsame Zentrifugation aus der Kultur entfernt, und die Überstandsfraktion wird zur weiteren Reinigung aufbewahrt.

35 Die Überstandsfraktion aus jedem Reinigungsverfahren wird einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Molekül entweder auf dem Chromatographieharz zurückgehalten wird, viele Verunreinigungen in der Probe jedoch nicht, oder die Verunreinigungen auf dem Harz zurückbleiben, die Probe  
40 hingegen nicht. Diese Chromatographieschritte können wenn nötig wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl geeigneter Chromatographieharze und ihrer wirksamsten Anwendung für ein bestimmtes zu reinigendes Molekül bewandert. Das  
45 gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

## 110

Im Fachgebiet ist ein breites Spektrum an Reinigungsverfahren bekannt, und das vorstehende Reinigungsverfahren soll nicht beschränkend sein. Diese Reinigungsverfahren sind zum Beispiel beschrieben in Bailey, J.E., & Ollis, D.F., Biochemical  
5 Engineering Fundamentals, McGraw-Hill: New York (1986).

Die Identität und Reinheit der isolierten Verbindungen kann durch Standardtechniken des Fachgebiets bestimmt werden. Dazu gehören Hochleistungs-Flüssigkeitschromatographie (HPLC), spektro-  
10 skopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, insbesondere Dünnschichtchromatographie und Flammenionisationsdetektion (IATROSCAN, Iatron, Tokio, Japan), NIRS, Enzymtest oder mikrobiologisch. Eine Übersicht über diese Analyseverfahren siehe in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140;  
15 Malakhova et al. (1996) Biotekhnologiya 11:27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Ed. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry  
20 and Molecular Biology, John Wiley and Sons; Fallon, A., et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Ed. 17.

## Äquivalente

25 Der Fachmann erkennt oder kann viele Äquivalente der hier beschriebenen erfindungsgemäßen spezifischen Ausführungsformen feststellen, indem er lediglich Routineexperimente verwendet. Diese Äquivalente sollen von den Patentansprüchen umfasst sein.

30

35

40

45

## Patentansprüche

1. Expressionskassette mit einer Struktur ausgewählt aus der  
5 Gruppe:
- a) L1 - Promotor - Strukturgen - L2,
- b) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2,  
10
- c) L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2 - L1 - Promotor - Strukturgen - L2,
- 15 wobei L1, L2, Promotor und Strukturgen die folgende Bedeutung hat:
- L1 = SEQ ID NO: 32 oder eine äquivalente Restriktionsschnittstellen enthaltende Sequenz,
- 20 L2 = unabhängig voneinander SEQ ID NO: 33, 34 oder 35 oder äquivalente Restriktionsschnittstellen enthaltende Sequenzen,
- Promotor = pflanzlicher Promotor
- Strukturgen = eine in Pflanzen exprimierbare  
25 Nukleinsäuresequenz.
2. Expressionskassette nach Anspruch 1, wobei das Strukturgen ein Biosynthesegen ist.
- 30 3. Expressionskassette nach Anspruch 1 oder 2, wobei das Strukturgen ein Biosynthesegen des Lipid- oder Fettsäurestoffwechsels ist.
4. Expressionskassette nach den Ansprüchen 1 bis 3, wobei das  
35 Strukturgen ein pflanzliches Gen ist.

40

Zeichn.

45

## 112

5. Expressionskassette nach den Ansprüche 1 bis 4, wobei das Gen eine Nukleinsäuresequenz ist, die für Proteine ausgewählt aus der Gruppe :
- 5 Acyl-CoA-Dehydrogenase(n), Acyl-ACP[= acyl carrier protein]-Desaturase(n), Acyl-ACP-Thioesterase(n), Fettsäure-Acyl-Transferase(n), Fettsäure-Synthase(n), Fettsäure-Hydroxylase(n), Acetyl-Coenzym A-Carboxylase(n), Acyl-Coenzym A-Oxidase(n), Fettsäure-Desaturase(n), Fettsäure-Acetylenasen, Lipoxxygenasen, Triacylglycerol-Lipasen,
- 10 Allenoxid-Synthasen, Hydroperoxid-Lyasen oder Fettsäure-Elongase(n), kodiert.
6. Expressionskassette nach den Ansprüche 1 bis 5, wobei das Gen eine Nukleinsäuresequenz ausgewählt aus der Gruppe ist:
- 15 a) einer Nukleinsäuresequenz mit der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Sequenz,
- 20 b) Nukleinsäuresequenzen, die aufgrund des degenerierten genetischen Codes durch Rückübersetzung der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen erhalten werden,
- 25 c) Derivate der in SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5 oder SEQ ID NO: 11 dargestellten Nukleinsäuresequenz, die für Polypeptide mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6 oder SEQ ID NO: 12 dargestellten Aminosäuresequenzen codieren und mindestens 50 % Homologie auf
- 30 Aminosäureebene aufweisen, ohne daß die enzymatische Wirkung der Polypeptide wesentlich reduziert ist.
7. Verwendung von Expressionskassetten gemäß den Ansprüchen 1 bis
- 35 6 in einem Verfahren zur Herstellung von Fettsäureestern mit einem erhöhtem Gehalt an mehrfach ungesättigten Fettsäuren mit mindestens zwei Doppelbindungen dadurch gekennzeichnet, daß man die Expressionskassette in einen Fettsäureester produzierenden Organismus einbringt, anzieht und die in dem Or-
- 40 ganismus enthaltenen Fettsäureester isoliert.
8. Verfahren nach Anspruch 7, wobei die durch das Verfahren hergestellten Fettsäureester mehrfach ungesättigte C<sub>18</sub>-, C<sub>20</sub>- oder C<sub>22</sub>-Fettsäuremoleküle mit mindestens zwei Doppelbindungen
- 45 im Fettsäureester enthalten.

## 113

9. Verfahren nach Anspruch 7 oder 8, wobei die C<sub>18</sub>-, C<sub>20</sub>- oder C<sub>22</sub>-Fettsäuremoleküle aus dem Organismus in Form eines Öls, Lipids isoliert werden.
- 5 10. Verfahren nach den Ansprüchen 7 bis 9, wobei der Organismus eine transgene Pflanze ist.
11. Verfahren nach den Ansprüchen 7 bis 10, wobei die Fettsäureester C<sub>18</sub>-, C<sub>20</sub>- oder C<sub>22</sub>-Fettsäuren mit drei, vier oder fünf  
10 Doppelbindungen im Fettsäureester enthalten.
12. Verfahren nach den Ansprüchen 7 bis 11, dadurch gekennzeichnet, daß man die in den Fettsäureestern enthaltenden mehrfach ungesättigten Fettsäuren freisetzt.
- 15 13. Vektor enthaltend eine Expressionskassette gemäß den Ansprüchen 1 bis 6.
14. Organismus enthaltend mindestens eine Expressionskassette gemäß den Ansprüchen 1 bis 6 oder mindestens einen Vektor gemäß  
20 Anspruch 13.
15. Organismus nach Anspruch 14, wobei es sich bei dem Organismus um eine Pflanze handelt.
- 25 16. Transgene Pflanze enthaltend eine funktionelle oder nicht funktionelle Nukleinsäuresequenz in einer Expressionskassette gemäß den Ansprüchen 1 bis 6.
- 30 17. Verwendung einer einer Expressionskassette gemäß den Ansprüchen 1 bis 6 zur Herstellung von transgenen Pflanzen.

35

40

45

**Figur 1:** Polypeptidvergleich der Codierregionen von Pp\_des 6 (obere Reihe) mit der EST-Sequenz von PT001078032R (untere Reihe)

```

398 WKPLVWMAVTELMMSGMLLGFVFLSHNGMEVYNSSKEFVSAQI-----VSTR 444
      W+      + + +      + L +F LSHN      + S+      +A +      V T
430 WRVFGNIMMGVAESLALAVLFSLSHN----FESADRDPTAPLKKTGEPVDWFKTQVETS 263
445 DIKGNIFNDWFTGGLNRQIEHHLFPTMPRHNLNKIAPRVEVFCKKHGLVY 494
      G      + FTGGLN Q+EHHLFP M      IAP+V      C KHG+ Y
262 CTYGGFLSGCFTGGLNFQVEHHLFPRMSSAWYXYIAPKVREICAKHGVHY 113

```

**Figur 2:** Polypeptidvergleich von Codierregionen der Ma\_des12 (obere Sequenz) mit PT001070010R

```

105 GVWVLAHECGHQSFSSTKTLNN 126
      G WVLAHECGH +FS +++L +
533 GFWVLAHECGHGAFSKNRSLQD 598

```

**Figur 2a:** Polypeptidvergleich von Codierregionen der Ma\_des12 (obere Sequenz) mit PT001072031R

```

117 SFSTSKTLNNTVGVWILHSMLLVPHYHSWRISHSKHH 151
      ++S S+T N+ VG+I+H LLVPY +W+ +H+KHH
465 AYSDSQTFNDVVGVFIVHQALLVPYFAWQYTHAKHH 569

```

**Figur 3:** Polypeptidvergleich von Codierregionen eines PCR Produktes aus Primerpaar F6a und R4a2 codierend für ein Desaturase Fragment (obere Reihe) aus *Phaodactylum* mit der Sequenz T36617 aus *Streptomyces coelicolor* (untere Reihe)

```

1  WWKNKHNGHHA VFN LHCSSAVAQDGD PDIDT M PLLAWSVQQAQSYRELQADGKDSGLVKF 60
      WW++KH HHA PN      +D DPDI      LL WS QA++      +GL +
114 WWQDKHTRHEANPN-----TEDLDPDIGP-DLLVWSPDQARAA-----TGLPRL 156
61  MIRNQSYFYFPILL LARLSWLNESFKCAFGLGAASENAAL ELKAKGLQYPLLEKAGILLH 120
      + R Q++ +FP+L L      E F      G A N L+ +A      L+ A +L H
157 LGRWQAFLEFP LLTL-----EGFNLHVASGRAMANRR LKRR-----LDGALLLAH 202
121 YAWMLTVSSGFRXXXXXXXXXXXXXXXXXCGFLLAIVFGLGHNGMATYNADARPDFWK LQ 180
      A LT      F      G L      F H GM      AD RPDF + Q
203 CAVYLTAL--FWVLPPGMAIAFLAVHQCLFGVYLGSAFAFNHKGMPILTADDRPDFLRRQ 260
181 VTTTRNVTGGHGFPQAFVDWFCGGLQYQVDHHLFPS 216
      V T+RNV GG      F D      GGL +Q++HHLFPS
261 VLTSRNVNGG-----LFTDLALGGLNHQIEHHLFPS 291

```

**Figur 4:** Polypeptidvergleich von Codierregionen aus Pp\_des6  
(obere Reihe) verglichen mit Pt\_des6 (untere Reihe)

```

51 KRLTSKKRVSESAAVQCISAEVQRNSSTQGTAEALAESVVKPTRRRSSQW 100
      . |   |.
1  .....MGKGGDARASKG 12

101 KKSTHPLS..EVAVHNKPSDCWIVVKNKVYDVSNFADHPPGGSVISTYFG 148
      :| || | | ||: |||||. |||||.|| |: |
13 STAARKISWQEVKTHASPEDAWIIHSNKVYDVSNW.HEHPPGGAIVFTHAG 61

149 RDGTDVFSSFHAASTWKILQDFYIGDV..ERVEPTPELL...KDFREMRA 193
      | ||:|..||| . :. ||||:. |   |: : | :||:|.
62 DDMFDIFAAPHAPGSQSLMKKFYIGELLPETTGKEPQQIAFEKGYRDLRS 111

194 LFLREQLFKSSKLYYVMKLLTNVAIFAASIAIICWSKTISAVLASACMMA 243
      : :|||. :|| | |. |||. ||. |:: :|   ||| |:
112 KLIMMGFMFSNKWFYVYKCLSNMAIWAAACALVFYSDFWVHLASAVMLG 161

244 LCFQQCGWLSHDFLHNQVFETRWNNEVVGIVIGNAVLGFSTGWWKEKHNL 293
      ||| |||. |||||. ||| | :. |   || . |:| ||| |||
162 TFFQQSGWLAHDFLHHQVFTKRKHGDLGGLFWGNLMQGYSVQWVKNKHNG 211

294 HHAAFN.ECDQTY.QPIDEDIDTLP LIAWS.....KDILATVENKTFL 334
      ||| || |   | | ||||:|:| ||   ::: | .. .
212 HHAVPNLHCSSAVAQDGDPDIDTMPLLAWSVQQAQSYRELQADGKDSGLV 261

335 R.ILQYQHLFFMGLLFFARGSWLFWSWR.....YTSTAVLSPVDR... 373
      : .. | |: :| || ||| |.: . | | :
262 KFMIRNQSYFYFPILLRLARLSWLNESFKCAFGLGAASENAALELKAKGLQ 311

374 ..LLEKGTVLFFHYFWFVG TAC.YLLPGWKPLVWMAVTELMS.GMLLGFVF 419
      |||| :| || | . . :. . . |   || || ||
312 YP LLEKAGILLHYAWMLTVSSGFGFRFSFAYTAFYFLTATASCGFLLAIVF 361

420 VLSHNGMEVYNSS..KEFVSAQIVSTRDIKG.....NIFNDWFTGGLNRQ 462
      | |||| ||. :| |: .||.: |   | ||| ||| |
362 GLGHNGMATYNADARPDFWKLVTTTRNVTGGHGFPAFVDWFCCGLQYQ 411

463 IEHHLFPTMPRHNLNKIAPRVEVFCKKHGLVYEDVSIATGTCKVLKALKE 512
      ::|||||.:|||| |   || |||. |. | : : || .|| |
412 VDHHLFPSLPRHNLAKTHALVESFCKEWGVQYHEADLVDGTMEVLHHLGS 461

513 VAEAAAQEHATTS.... 525
      ||           |
462 VAGEFVVDFVRDGPAM. 477

```

**Figur 5:** Polypeptidvergleich von Codierregionen aus Pp\_des6 (obere Reihe) verglichen mit Pt\_des5 (untere Reihe)

```

51 KRLTSKKRVSESAAVQCISAEVQRNSSSTQGTAEALAESVVKPTRRSSQW 100
      : | | .|...
1  .....MAPDADKLRQRQTAV 16
      . . . . .
101 KKSTHPLSEVAVHNKPSDC.....WIVVKNKVYDVSNFADEHPGGSVIS 144
    | | . : : : | | . | : || | |
17 AK..HNAATISTQERLCSLSSLKGEEVCIDGIIYDLQSF..DHPGGETIK 62
      . . . . .
145 TYFGRDGTDFVSSFHAASWTKILQDF.YIGDVERVEPTPELLKDF.REM. 191
    : | | | : | | | | : | | . : | ||.
63 MFGGNDVTVQYKMIHPYHTEKHLEKMKRVGKVTDVFCYKFDTEFEREIK 112
      . . . . .
192 RALFLREQLFKS.SKLYYVMKLLTNVAIFAASIAIICWSKT.ISAVLASA 239
    | . | : : : || | | | | . | | |
113 REVFKIVRRGKDFGTGLGWFFRAFCYIAIF..FYLQYHWVTGTGSWLLAVA 160
      . . . . .
240 CMMALCFQQCGWLSHDFLHNQVFETRNLNEVVGIVIGNAVLGFSTGWWE 289
    .. . || | . | : : | : | | | . |
161 YGISQAMIGMN.VQHDANHGATSKRPWVNDMLG..LGADFIGGSKWLWQE 207
      . . . . .
290 KHNLHHAAPNECDQTYQPIDEDIDTLPLIAWSKDILATVENKTFLRLIQY 339
    . | || | | : | : | : : . | : | . :
208 QHWTTHAYTNHAEM..DP..DSFGAEPMLLFN.DYPLDHPARTWLH..RF 250
      . . . . .
340 QHLFFMGLLEFFARGSWLFWSWR.....YTSTAVLS...PVDRLLEKGTVL 381
    | | : | | | | | . || . |
251 QAFFYMPVL...AGYWLSAVFNPQILDLOQRGALSVGIRLDNAFIHSRRK 297
      . . . . .
382 FHYFW...FVG....TACYLLPG....WKPLVWMAVTELMGMLLGFVFW 420
    : || : : | | | : . . : | . |
298 YAVFWRAVYIAVNVIAPFYTNSGLEWSWRVFGNIMLMGVAESLALAVLFS 347
      . . . . .
421 LSHN.....GMEVYNSSKEFVSAQIVSTRDIKGNIFNDWFTGGLN 460
    |||| :. :. | | | | . |||||
348 LSHNFESADRDPTAPLKKTGEPVDWFKTQ.VETSCTYGGFLSGCFTGGLN 396
      . . . . .
461 RQIBHHLFPTMPRHNLNKIAPRVEVFCKKHGLVYEDVS.IATGTCKVLKA 509
    | : |||| | ||| : | |||. | . :
397 FQVEHHLFPRMSSAWYPYIAPKVRERICAKHGWHYAYYPWIHQNFLSTVRY 446
      . . . . .
510 LKEVAEAAA.EQHATTS..... 525
    : | | | . |
447 MHAAGTGANWROMARENPLTGRA. 469

```

**Figur 6:** Polypeptidvergleich von Codierregionen der  $\Delta$ -12-Desaturase aus *Mortierella alpina* (Ma\_des12) obere Reihe mit der homologen Sequenz aus *Phaeodactylum tricornutum* (Pt\_des12) in der unteren Reihe

```

40 KEIRECIPAHCFERSGLRGLCHVAIDL TWASLL..FLAATQIDKFE..NP 85
   |::| || ||| : |::: | :| . | : : ||
107 KDLRAVIPKDCFEPTAKSLGYLSVS.TMGTILCSVVGANLLSVLDPSNP 155

      .
      .
86 LIRYLAWPVYWIMQGIVCTGVWVLAHECGHQSFSKTLNNTVGWILHSM 135
   | : | | . | | |.||||||| |.|| :. | . ||:|:|.
156 L.TWPLWAAYGAVTGTVAMGLWVLAHECGHGAFSKNRSIQDAVGYYIHSI 204

      .
      .
136 LLVZYHSWRISHSKHHKATGHMTKDQVFVPKTRSQVGLPPKENAAA VQE 185
   :||| ||. ||. ||. | || : || . | | . | |
205 MLVPYFSWQRSHAVHHQYTNHMLGETHV PDRADKEG....EKSLALRQF 250

      .
      .
186 EDMSVHLDEEAPIVTLFWMVIQFLFGWPAYLIMNASGQDYGRWTS HFHTY 235
   | | . : : |||||: . |. |. ||:
251 MLDSFGKDKGMKAYGGLQSFHLIVGWPAYLLIGATGGPDRGMTNHFYP. 299

      .
      .
236 SPIFEP....RNFF.....DIIISDLGVLAALGALIYASMQLSLLTVTK 275
   .|: | : | : ||:|: |.|||| . | |
300 NPLSTPTQPKKELFPGNWKEKVYQSDIGIAAVVGALIAWTATSG LAPVMA 349

      .
      .
276 YYIVPYLFVNFWLVLTITFLQHTDPKLPHYREGAWN FQRGALCTVDRSFGK 325
   | | : :| |||| |.|||| |.||: || :|| |:| : |
350 LYGGPLIVINAWLVLYTWLQHTDITDVPHFSSDNHNFVKALHTIDR PYDK 399

      .
      .
326 .....FLDHMFHGIVHTHVAHHLFSQMPFYHAE EATYHLKKLLGEYYVYD 370
   :| : | | ||||| |. | | :| || :| | |. ||
400 LDPWGIIDFLHHKIGTTHVAHHFDSTIPHYKAQIATDAIKAKFPEVYLYD 449

      .
      .
371 PSPIVVAVWRSFREC RFVEDQGDVVFFK 398
   |.|| |.|| : | || .| .|
450 PTPIPQAMWRVAKGCTAVEQRGDAWVWK 477

```

**Figur 7:** Polypeptidvergleich von Codierregionen der  $\Delta$ -12-Desaturase aus *Mortierella alpina* (Ma\_des12) obere Reihe mit der homologen Sequenz aus *Phaeodactylum tri-cornutum* Klon PT001072031R (Pt\_des12.2) in der unteren Reihe.

```

22 NSAKPAFERNYQLPEFTIKEIRECIPAHCFERSGLRGLCHVAIDL TWASL 71
   .| | . . :|| | :|: || ||:| | :. || |.
33 SSYNPLAKDSPELP..TKGQIKAVIPKECFORSAFWSTFYLMRDLAMAAA 80

72 LFLAATQIDKFENP.....LIRYLAWPVWIMQGIVCTGVWVLAHECGH 115
   .|: : | | | | | | | | | | | | | | | |
81 FCYGTSQVLSTDLPPQDATLILPWALGWGVYAFWMGTILTGPWVVAHECGH 130

116 QSFSTSKTLNNTVGWILHSMLLVPYHSWRISHSKHHKATGHMTKDQVFVP 165
   .:| |. |. |||. | |||| |. .|. |||: | |: : ||
131 GAYSDSQTFNDVVGFIHVHQALLVPYFAWQYTHAKHHRRTNHLVDGESHPV 180

166 KTRSQVGLPP..KENAAAVQEEDMSVHLDEEAPIVTLFWMVIQFLFGWP 213
   | || | . |. | | | : | | | | ||
181 STAKDNLGPHNERNNSFYAAWHEAMG....DGAFVAVFQVWS..HLFVGWP 224

214 AYLI.MNASGQ..DYGRW.....TSHFHTYSPIFEPRNFFDIIISDLG 253
   || : ..|. | | | | | | | | | | | | | | | |
225 LYLAGLASTGKLAHEGWLEERNAIADHFRPSSPMFPKIRAKIALSSAT 274

254 VLAALGALIYASMQLSLLTVTKYYIVPYLFVNFVLVLITFLQHTDPKLP 303
   || | |:| |. | | :| || ||| |||| |. ||||| | :||
275 ELAVLAGLLYVGTQVGHLPVLLWYWGPTYTFVNAWLVLTYLQHTDPSIPH 324

304 YREGAWNFORGALCTVDRSFGKFLDHMFHGIVHTHVAHHLFSQMPFYHAE 353
   | || | . :|| | :|| :| | | | | | | | | | :|||. |
325 YGEGETWVKGALSTIDRDYGIF.DFFHHTIGSTHVHHLFHEMPWYNAG 373

354 EATYHLKLLGE..YYVYDPSPIVVAVVWRSFRECRFVEDQGDVVFFK 398
   || .|. | | |||. | |. | | | :|| | :||
374 IATQKVKEFLEPQGLYNDPTPWYKAMWRIARTCHYVESNEGVOYFK 420

```

## SEQUENZPROTOKOLL

&lt;110&gt; BASF Plant Science GmbH

<120> Verfahren zur Expression von Biosynthesegenen in  
pflanzlichen Samen unter Verwendung von neuen  
multiplene Expressionskonstrukten

&lt;130&gt; 2000\_904

&lt;140&gt; 2000\_904

&lt;141&gt; 2000-12-22

&lt;160&gt; 35

&lt;170&gt; PatentIn Vers. 2.0

&lt;210&gt; 1

&lt;211&gt; 1652

&lt;212&gt; DNA

&lt;213&gt; Phaeodactylum tricornutum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (115)..(1524)

&lt;400&gt; 1

gacccaacaa acccaacaat cccaacaatc ccattaacag gaattgggtt tcgttgagtc 60

aataattgct agaattccaaa cagacagaca gagaccaacc gcattctatta caga atg 117  
Met  
1gct ccg gat gcg gat aag ctt cga caa cgc cag acg act gcg gta gcg 165  
Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val Ala  
5 10 15aag cac aat gct gct acc ata tcg acg cag gaa cgc ctt tgc agt ctg 213  
Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser Leu  
20 25 30tgt tcg ctc aaa ggc gaa gaa gtc tgc atc gac gga atc atc tat gac 261  
Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr Asp  
35 40 45ctc caa tca ttc gat cat ccc ggg ggt gaa acg atc aaa atg ttt ggt 309  
Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe Gly  
50 55 60 65ggc aac gat gtc act gta cag tac aag atg att cac ccg tac cat acc 357  
Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His Thr  
70 75 80gag aag cat ttg gaa aag atg aag cgt gtc ggc aag gtg acg gat ttc 405  
Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp Phe  
85 90 95

gtc tgc gag tac aag ttc gat acc gaa ttt gaa cgc gaa atc aaa cga 453

Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys Arg	
100 105 110	
gaa gtc ttc aag att gtg cga cga ggc aag gat ttc ggt act ttg gga	501
Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu Gly	
115 120 125	
tgg ttc ttc cgt gcg ttt tgc tac att gcc att ttc ttc tac ctg cag	549
Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu Gln	
130 135 140 145	
tac cat tgg gtc acc acg gga acc tct tgg ctg ctg gcc gtg gcc tac	597
Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala Tyr	
150 155 160	
gga atc tcc caa gcg atg att ggc atg aat gtc cag cac gat gcc aac	645
Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala Asn	
165 170 175	
cac ggg gcc acc tcc aag cgt ccc tgg gtc aac gac atg cta ggc ctc	693
His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly Leu	
180 185 190	
ggt gcg gat ttt att ggt ggt tcc aag tgg ctc tgg cag gaa caa cac	741
Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln His	
195 200 205	
tgg acc cac cac gct tac acc aat cac gcc gag atg gat ccc gat agc	789
Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp Ser	
210 215 220 225	
ttt ggt gcc gaa cca atg ctc cta ttc aac gac tat ccc ttg gat cat	837
Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp His	
230 235 240	
ccc gct cgt acc tgg cta cat cgc ttt caa gca ttc ttt tac atg ccc	885
Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met Pro	
245 250 255	
gtc ttg gct gga tac tgg ttg tcc gct gtc ttc aat cca caa att ctt	933
Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile Leu	
260 265 270	
gac ctc cag caa cgc ggc gca ctt tcc gtc ggt atc cgt ctc gac aac	981
Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp Asn	
275 280 285	
gct ttc att cac tcg cga cgc aag tat gcg gtt ttc tgg cgg gct gtg	1029
Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala Val	
290 295 300 305	
tac att gcg gtg aac gtg att gct ccg ttt tac acā aac tcc ggc ctc	1077
Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly Leu	
310 315 320	
gaa tgg tcc tgg cgt gtc ttt gga aac atc atg ctc atg ggt gtg gcg	1125
Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val Ala	
325 330 335	

gaa tcg ctc gcg ctg gcg gtc ctg ttt tcg ttg tcg cac aat ttc gaa 1173  
 Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe Glu  
 340 345 350

tcc gcg gat cgc gat ccg acc gcc cca ctg aaa aag acg gga gaa cca 1221  
 Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu Pro  
 355 360 365

gtc gac tgg ttc aag aca cag gtc gaa act tcc tgc act tac ggt gga 1269  
 Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly Gly  
 370 375 380 385

ttc ctt tcc ggt tgc ttc acg gga ggt ctc aac ttt cag gtt gaa cac 1317  
 Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu His  
 390 395 400

cac ttg ttc cca cgc atg agc agc gct tgg tat ccc tac att gcc ccc 1365  
 His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro  
 405 410 415

aag gtc cgc gaa att tgc gcc aaa cac gcc gtc cac tac gcc tac tac 1413  
 Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr Tyr  
 420 425 430

ccg tgg atc cac caa aac ttt ctc tcc acc gtc cgc tac atg cac gcg 1461  
 Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His Ala  
 435 440 445

gcc ggg acc ggt gcc aac tgg cgc cag atg gcc aga gaa aat ccc ttg 1509  
 Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro Leu  
 450 455 460 465

acc gga cgg gcg taa aagtacacga cagcaccaaa ggtggcgtat ggtgatctct 1564  
 Thr Gly Arg Ala  
 470

agaaaacaga catagcctac tggaaatata gacgtccaaa caataatttt aaagactatt 1624

tttctgcgta aaaaaaaaaa aaaaaaaa 1652

&lt;210&gt; 2

&lt;211&gt; 469

&lt;212&gt; PRT

<213> *Phaeodactylum tricornutum*

&lt;400&gt; 2

Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Gln Thr Thr Ala Val  
 1 5 10 15

Ala Lys His Asn Ala Ala Thr Ile Ser Thr Gln Glu Arg Leu Cys Ser  
 20 25 30

Leu Ser Ser Leu Lys Gly Glu Glu Val Cys Ile Asp Gly Ile Ile Tyr  
 35 40 45

Asp Leu Gln Ser Phe Asp His Pro Gly Gly Glu Thr Ile Lys Met Phe

50	55	60
Gly Gly Asn Asp Val Thr Val Gln Tyr Lys Met Ile His Pro Tyr His		
65	70	75 80
Thr Glu Lys His Leu Glu Lys Met Lys Arg Val Gly Lys Val Thr Asp		
	85	90 95
Phe Val Cys Glu Tyr Lys Phe Asp Thr Glu Phe Glu Arg Glu Ile Lys		
	100	105 110
Arg Glu Val Phe Lys Ile Val Arg Arg Gly Lys Asp Phe Gly Thr Leu		
	115	120 125
Gly Trp Phe Phe Arg Ala Phe Cys Tyr Ile Ala Ile Phe Phe Tyr Leu		
	130	135 140
Gln Tyr His Trp Val Thr Thr Gly Thr Ser Trp Leu Leu Ala Val Ala		
	145	150 155 160
Tyr Gly Ile Ser Gln Ala Met Ile Gly Met Asn Val Gln His Asp Ala		
	165	170 175
Asn His Gly Ala Thr Ser Lys Arg Pro Trp Val Asn Asp Met Leu Gly		
	180	185 190
Leu Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Gln Glu Gln		
	195	200 205
His Trp Thr His His Ala Tyr Thr Asn His Ala Glu Met Asp Pro Asp		
	210	215 220
Ser Phe Gly Ala Glu Pro Met Leu Leu Phe Asn Asp Tyr Pro Leu Asp		
	225	230 235 240
His Pro Ala Arg Thr Trp Leu His Arg Phe Gln Ala Phe Phe Tyr Met		
	245	250 255
Pro Val Leu Ala Gly Tyr Trp Leu Ser Ala Val Phe Asn Pro Gln Ile		
	260	265 270
Leu Asp Leu Gln Gln Arg Gly Ala Leu Ser Val Gly Ile Arg Leu Asp		
	275	280 285
Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala		
	290	295 300
Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly		
	305	310 315 320
Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val		
	325	330 335
Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe		
	340	345 350
Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu		
	355	360 365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly  
370 375 380

Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu  
385 390 395 400

His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala  
405 410 415

Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr  
420 425 430

Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His  
435 440 445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro  
450 455 460

Leu Thr Gly Arg Ala  
465

<210> 3

<211> 1434

<212> DNA

<213> Phaeodactylum tricornutum

<220>

<221> CDS

<222> (1)..(1434)

<400> 3

atg ggc aaa gga ggg gac gct cgg gcc tcg aag ggc tca acg gcg gct 48  
Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala  
1 5 10 15

cgc aag atc agt tgg cag gaa gtc aag acc cac gcg tct ccg gag gac 96  
Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp  
20 25 30

gcc tgg atc att cac tcc aat aag gtc tac gac gtg tcc aac tgg cac 144  
Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His  
35 40 45

gaa cat ccc gga ggc gcc gtc att ttc acg cac gcc ggt gac gac atg 192  
Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met  
50 55 60

acg gac att ttc gct gcc ttt cac gca ccc gga tcg cag tcg ctc atg 240  
Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met  
65 70 75 80

aag aag ttc tac att ggc gaa ttg ctc ccg gaa acc acc ggc aag gag 288  
Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu  
85 90 95

ccg cag caa atc gcc ttt gaa aag ggc tac cgc gat ctg cgc tcc aaa 336

Pro	Gln	Gln	Ile	Ala	Phe	Glu	Lys	Gly	Tyr	Arg	Asp	Leu	Arg	Ser	Lys	
			100					105					110			
ctc	atc	atg	atg	ggc	atg	ttc	aag	tcc	aac	aag	tgg	ttc	tac	gtc	tac	384
Leu	Ile	Met	Met	Gly	Met	Phe	Lys	Ser	Asn	Lys	Trp	Phe	Tyr	Val	Tyr	
		115					120					125				
aag	tgc	ctc	agc	aac	atg	gcc	att	tgg	gcc	gcc	gcc	tgt	gct	ctc	gtc	432
Lys	Cys	Leu	Ser	Asn	Met	Ala	Ile	Trp	Ala	Ala	Ala	Cys	Ala	Leu	Val	
		130				135					140					
ttt	tac	tcg	gac	cgc	ttc	tgg	gta	cac	ctg	gcc	agc	gcc	gtc	atg	ctg	480
Phe	Tyr	Ser	Asp	Arg	Phe	Trp	Val	His	Leu	Ala	Ser	Ala	Val	Met	Leu	
145					150					155					160	
gga	aca	ttc	ttt	cag	cag	tcg	gga	tgg	ttg	gca	cac	gac	ttt	ctg	cac	528
Gly	Thr	Phe	Phe	Gln	Gln	Ser	Gly	Trp	Leu	Ala	His	Asp	Phe	Leu	His	
				165					170					175		
cac	cag	gtc	ttc	acc	aag	cgc	aag	cac	ggg	gat	ctc	gga	gga	ctc	ttt	576
His	Gln	Val	Phe	Thr	Lys	Arg	Lys	His	Gly	Asp	Leu	Gly	Gly	Leu	Phe	
			180					185						190		
tgg	ggg	aac	ctc	atg	cag	ggt	tac	tcc	gta	cag	tgg	tgg	aaa	aac	aag	624
Trp	Gly	Asn	Leu	Met	Gln	Gly	Tyr	Ser	Val	Gln	Trp	Trp	Lys	Asn	Lys	
		195					200					205				
cac	aac	gga	cac	cac	gcc	gtc	ccc	aac	ctc	cac	tgc	tcc	tcc	gca	gtc	672
His	Asn	Gly	His	His	Ala	Val	Pro	Asn	Leu	His	Cys	Ser	Ser	Ala	Val	
		210				215					220					
gcg	caa	gat	ggg	gac	ccg	gac	atc	gat	acc	atg	ccc	ctt	ctc	gcc	tgg	720
Ala	Gln	Asp	Gly	Asp	Pro	Asp	Ile	Asp	Thr	Met	Pro	Leu	Leu	Ala	Trp	
225					230					235					240	
tcc	gtc	cag	caa	gcc	cag	tct	tac	cgg	gaa	ctc	caa	gcc	gac	gga	aag	768
Ser	Val	Gln	Gln	Ala	Gln	Ser	Tyr	Arg	Glu	Leu	Gln	Ala	Asp	Gly	Lys	
				245					250					255		
gat	tcg	ggt	ttg	gtc	aag	ttc	atg	atc	cgt	aac	caa	tcc	tac	ttt	tac	816
Asp	Ser	Gly	Leu	Val	Lys	Phe	Met	Ile	Arg	Asn	Gln	Ser	Tyr	Phe	Tyr	
			260					265						270		
ttt	ccc	atc	ttg	ttg	ctc	gcc	cgc	ctg	tcg	tgg	ttg	aac	gag	tcc	ttc	864
Phe	Pro	Ile	Leu	Leu	Leu	Ala	Arg	Leu	Ser	Trp	Leu	Asn	Glu	Ser	Phe	
		275					280						285			
aag	tgc	gcc	ttt	ggg	ctt	gga	gct	gcg	tcg	gag	aac	gct	gct	ctc	gaa	912
Lys	Cys	Ala	Phe	Gly	Leu	Gly	Ala	Ala	Ser	Glu	Asn	Ala	Ala	Leu	Glu	
		290				295					300					
ctc	aag	gcc	aag	ggt	ctt	cag	tac	ccc	ctt	ttg	gaa	aag	gct	ggc	atc	960
Leu	Lys	Ala	Lys	Gly	Leu	Gln	Tyr	Pro	Leu	Leu	Glu	Lys	Ala	Gly	Ile	
305					310					315					320	
ctg	ctg	cac	tac	gct	tgg	atg	ctt	aca	gtt	tcg	tcc	ggc	ttt	gga	cgc	1008
Leu	Leu	His	Tyr	Ala	Trp	Met	Leu	Thr	Val	Ser	Ser	Gly	Phe	Gly	Arg	
				325					330					335		

ttc tgc ttc gcg tac acc gca ttt tac ttt cta acc gcg acc gcg tcc 1056  
 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser  
 340 345 350  
 tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc ggc cac aac ggc atg 1104  
 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met  
 355 360 365  
 gcc acc tac aat gcc gac gcc cgt ccg gac ttc tgg aag ctc caa gtc 1152  
 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val  
 370 375 380  
 acc acg act cgc aac gtc acg ggc gga cac ggt ttc ccc caa gcc ttt 1200  
 Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe  
 385 390 395 400  
 gtc gac tgg ttc tgt ggt ggc ctc cag tac caa gtc gac cac cac tta 1248  
 Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu  
 405 410 415  
 ttc ccc agc ctg ccc cga cac aat ctg gcc aag aca cac gca ctg gtc 1296  
 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val  
 420 425 430  
 gaa tgc ttc tgc aag gag tgg ggt gtc cag tac cac gaa gcc gac ctt 1344  
 Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu  
 435 440 445  
 gtg gac ggg acc atg gaa gtc ttg cac cat ttg ggc agc gtg gcc ggc 1392  
 Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly  
 450 455 460  
 gaa ttc gtc gtg gat ttt gta cgc gat gga ccc gcc atg taa 1434  
 Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
 465 470 475

<210> 4

<211> 477

<212> PRT

<213> *Phaeodactylum tricornutum*

<400> 4

Met Gly Lys Gly Gly Asp Ala Arg Ala Ser Lys Gly Ser Thr Ala Ala  
 1 5 10 15

Arg Lys Ile Ser Trp Gln Glu Val Lys Thr His Ala Ser Pro Glu Asp  
 20 25 30

Ala Trp Ile Ile His Ser Asn Lys Val Tyr Asp Val Ser Asn Trp His  
 35 40 45

Glu His Pro Gly Gly Ala Val Ile Phe Thr His Ala Gly Asp Asp Met  
 50 55 60

Thr Asp Ile Phe Ala Ala Phe His Ala Pro Gly Ser Gln Ser Leu Met  
 65 70 75 80

Lys Lys Phe Tyr Ile Gly Glu Leu Leu Pro Glu Thr Thr Gly Lys Glu  
 85 90 95  
 Pro Gln Gln Ile Ala Phe Glu Lys Gly Tyr Arg Asp Leu Arg Ser Lys  
 100 105 110  
 Leu Ile Met Met Gly Met Phe Lys Ser Asn Lys Trp Phe Tyr Val Tyr  
 115 120 125  
 Lys Cys Leu Ser Asn Met Ala Ile Trp Ala Ala Ala Cys Ala Leu Val  
 130 135 140  
 Phe Tyr Ser Asp Arg Phe Trp Val His Leu Ala Ser Ala Val Met Leu  
 145 150 155 160  
 Gly Thr Phe Phe Gln Gln Ser Gly Trp Leu Ala His Asp Phe Leu His  
 165 170 175  
 His Gln Val Phe Thr Lys Arg Lys His Gly Asp Leu Gly Gly Leu Phe  
 180 185 190  
 Trp Gly Asn Leu Met Gln Gly Tyr Ser Val Gln Trp Trp Lys Asn Lys  
 195 200 205  
 His Asn Gly His His Ala Val Pro Asn Leu His Cys Ser Ser Ala Val  
 210 215 220  
 Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met Pro Leu Leu Ala Trp  
 225 230 235 240  
 Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu Gln Ala Asp Gly Lys  
 245 250 255  
 Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn Gln Ser Tyr Phe Tyr  
 260 265 270  
 Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp Leu Asn Glu Ser Phe  
 275 280 285  
 Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu Asn Ala Ala Leu Glu  
 290 295 300  
 Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu Glu Lys Ala Gly Ile  
 305 310 315 320  
 Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser Ser Gly Phe Gly Arg  
 325 330 335  
 Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu Thr Ala Thr Ala Ser  
 340 345 350  
 Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu Gly His Asn Gly Met  
 355 360 365  
 Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe Trp Lys Leu Gln Val  
 370 375 380

Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly Phe Pro Gln Ala Phe  
385 390 395 400

Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln Val Asp His His Leu  
405 410 415

Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Thr His Ala Leu Val  
420 425 430

Glu Ser Phe Cys Lys Glu Trp Gly Val Gln Tyr His Glu Ala Asp Leu  
435 440 445

Val Asp Gly Thr Met Glu Val Leu His His Leu Gly Ser Val Ala Gly  
450 455 460

Glu Phe Val Val Asp Phe Val Arg Asp Gly Pro Ala Met  
465 470 475

<210> 5

<211> 1651

<212> DNA

<213> *Phaeodactylum tricornutum*

<220>

<221> CDS

<222> (67)..(1554)

<400> 5

gaagaaggaa catataaaag taagccatct cctcggcacc atctaaagac ctaatatcta 60

ctcgtc atg gtt cgc ttt tca aca gcc gct cta ctt tct ctg tcg aca 108  
Met Val Arg Phe Ser Thr Ala Ala Leu Ser Leu Ser Thr  
1 5 10

ttg aca act tca tgt att ggt gcc ttc cag ctg tct tcg cca gca caa 156  
Leu Thr Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln  
15 20 25 30

ctt ccg aca agt agg ctt cgt cgg cat acg aac acg gcg ccg ctt tcg 204  
Leu Pro Thr Ser Arg Leu Arg Arg His Thr Asn Thr Ala Pro Leu Ser  
35 40 45

gcc gtg gcc gtc gac tcc ggt tct tcc gat ccg gcc ttg gta ggc aac 252  
Ala Val Ala Val Asp Ser Gly Ser Ser Asp Pro Ala Leu Val Gly Asn  
50 55 60

ctc ccc ctt ccc aac aac aat gat aat gag gac aag aac cgt aga atg 300  
Leu Pro Leu Pro Asn Asn Asn Asp Asn Glu Asp Lys Asn Arg Arg Met  
65 70 75

cca atg atg gac ttg aaa ggt att gct ctg tct ggt ctc aaa ggg caa 348  
Pro Met Met Asp Leu Lys Gly Ile Ala Leu Ser Gly Leu Lys Gly Gln  
80 85 90

gct ctt tcc gtc cga gcg gaa gat ttt cct cag gcg aaa gac ttg cgt 396  
Ala Leu Ser Val Arg Ala Glu Asp Phe Pro Gln Ala Lys Asp Leu Arg  
95 100 105 110

gcc gtc att ccg aaa gat tgc ttc gaa ccc gac acg gcc aaa tcg ttg	444
Ala Val Ile Pro Lys Asp Cys Phe Glu Pro Asp Thr Ala Lys Ser Leu	
115 120 125	
gga tat ctt tcc gtt tca act atg ggg aca att ctc tgc tcc gtc gtc	492
Gly Tyr Leu Ser Val Ser Thr Met Gly Thr Ile Leu Cys Ser Val Val	
130 135 140	
ggc gcg aac ctc ctt agt gtg ctc gat ccc tcc aat cca tta acc tgg	540
Gly Ala Asn Leu Leu Ser Val Leu Asp Pro Ser Asn Pro Leu Thr Trp	
145 150 155	
cct ctc tgg gcg gcc tac ggt gcc gtc acg ggg acg gtc gcc atg ggg	588
Pro Leu Trp Ala Ala Tyr Gly Ala Val Thr Gly Thr Val Ala Met Gly	
160 165 170	
ctt tgg gtg ctg gcc cac gaa tgc gga cac ggc gcc ttt tcc aaa aac	636
Leu Trp Val Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys Asn	
175 180 185 190	
cga tcc ctc cag gat gcc gtg ggg tac att atc cat tcc atc atg ctg	684
Arg Ser Leu Gln Asp Ala Val Gly Tyr Ile Ile His Ser Ile Met Leu	
195 200 205	
gtg cca tac ttt agt tgg cag cga tcg cat gcc gtg cat cac cag tat	732
Val Pro Tyr Phe Ser Trp Gln Arg Ser His Ala Val His His Gln Tyr	
210 215 220	
acc aat cat atg gaa ctg ggg gaa aca cac gtt cct gat cga gcc gat	780
Thr Asn His Met Glu Leu Gly Glu Thr His Val Pro Asp Arg Ala Asp	
225 230 235	
aag gag ggc gag aag agc ctg gcg ctc cgc cag ttc atg ttg gat tcc	828
Lys Glu Gly Glu Lys Ser Leu Ala Leu Arg Gln Phe Met Leu Asp Ser	
240 245 250	
ttt ggt aaa gac aag ggc atg aaa gca tac gga ggc ctc cag tcg ttt	876
Phe Gly Lys Asp Lys Gly Met Lys Ala Tyr Gly Gly Leu Gln Ser Phe	
255 260 265 270	
ttg cat ctc atc gtg gga tgg cca gcc tac ctc ctg atc ggt gcg acc	924
Leu His Leu Ile Val Gly Trp Pro Ala Tyr Leu Leu Ile Gly Ala Thr	
275 280 285	
ggt gga ccc gac cgt ggt atg acc aac cat ttt tat ccc aac cct ttg	972
Gly Gly Pro Asp Arg Gly Met Thr Asn His Phe Tyr Pro Asn Pro Leu	
290 295 300	
tcg acg cca aca cag ccc aag aaa gaa ctt ttc cct ggg aac tgg aaa	1020
Ser Thr Pro Thr Gln Pro Lys Lys Glu Leu Phe Pro Gly Asn Trp Lys	
305 310 315	
gaa aag gtc tac cag tca gat att gga atc gcc gcc gtt gtc ggc gcc	1068
Glu Lys Val Tyr Gln Ser Asp Ile Gly Ile Ala Ala Val Val Gly Ala	
320 325 330	
ctc att gct tgg acc gcc act tcg ggt cta gcc ccc gtc atg gcc ttg	1116

Leu Ile Ala Trp Thr Ala Thr Ser Gly Leu Ala Pro Val Met Ala Leu  
 335 340 345 350  
  
 tac ggt ggt ccc ttg atc gtc att aat gcc tgg ctg gta ctg tac acg 1164  
 Tyr Gly Gly Pro Leu Ile Val Ile Asn Ala Trp Leu Val Leu Tyr Thr  
 355 360 365  
  
 tgg ttg caa cat aca gat acc gat gtt ccg cac ttt tcc tcc gac aac 1212  
 Trp Leu Gln His Thr Asp Thr Asp Val Pro His Phe Ser Ser Asp Asn  
 370 375 380  
  
 cac aac ttt gtc aag ggc gca ctg cat acg atc gat cgt ccc tac gac 1260  
 His Asn Phe Val Lys Gly Ala Leu His Thr Ile Asp Arg Pro Tyr Asp  
 385 390 395  
  
 aaa ctt gat ccc tgg gga atc ata gac ttt ctg cac cac aag att gga 1308  
 Lys Leu Asp Pro Trp Gly Ile Ile Asp Phe Leu His His Lys Ile Gly  
 400 405 410  
  
 aca acg cat gtg gca cac cat ttt gac agt act atc ccc cac tat aag 1356  
 Thr Thr His Val Ala His His Phe Asp Ser Thr Ile Pro His Tyr Lys  
 415 420 425 430  
  
 gct cag att gct acc gat gcc atc aaa gcc aag ttt cca gaa gtg tac 1404  
 Ala Gln Ile Ala Thr Asp Ala Ile Lys Ala Lys Phe Pro Glu Val Tyr  
 435 440 445  
  
 ctc tat gac ccg aca cca att cca caa gcc atg tgg cgc gtc gcc aag 1452  
 Leu Tyr Asp Pro Thr Pro Ile Pro Gln Ala Met Trp Arg Val Ala Lys  
 450 455 460  
  
 gga tgt act gca gta gag caa cgc ggt gac gcc tgg gtg tgg aaa aac 1500  
 Gly Cys Thr Ala Val Glu Gln Arg Gly Asp Ala Trp Val Trp Lys Asn  
 465 470 475  
  
 gaa gga ata gaa gat ttg gtg gaa cat cgt caa agc aaa tta tcg agc 1548  
 Glu Gly Ile Glu Asp Leu Val Glu His Arg Gln Ser Lys Leu Ser Ser  
 480 485 490  
  
 gaa taa agcaacatat cgctttatgg aagaacaaac gtccattgtg taaaaccctg 1604  
 Glu  
 495  
  
 ataatttcaa tattgtgttt tgttttaaaa aaaaaaaaaa aaaaaaa 1651

&lt;210&gt; 6

&lt;211&gt; 495

&lt;212&gt; PRT

&lt;213&gt; Phaeodactylum tricornutum

&lt;400&gt; 6

Met Val Arg Phe Ser Thr Ala Ala Leu Leu Ser Leu Ser Thr Leu Thr  
 1 5 10 15

Thr Ser Cys Ile Gly Ala Phe Gln Leu Ser Ser Pro Ala Gln Leu Pro  
 20 25 30

Thr Ser Arg Leu Arg Arg His Thr Asn Thr Ala Pro Leu Ser Ala Val  
 35 40 45  
 Ala Val Asp Ser Gly Ser Ser Asp Pro Ala Leu Val Gly Asn Leu Pro  
 50 55 60  
 Leu Pro Asn Asn Asn Asp Asn Glu Asp Lys Asn Arg Arg Met Pro Met  
 65 70 75 80  
 Met Asp Leu Lys Gly Ile Ala Leu Ser Gly Leu Lys Gly Gln Ala Leu  
 85 90 95  
 Ser Val Arg Ala Glu Asp Phe Pro Gln Ala Lys Asp Leu Arg Ala Val  
 100 105 110  
 Ile Pro Lys Asp Cys Phe Glu Pro Asp Thr Ala Lys Ser Leu Gly Tyr  
 115 120 125  
 Leu Ser Val Ser Thr Met Gly Thr Ile Leu Cys Ser Val Val Gly Ala  
 130 135 140  
 Asn Leu Leu Ser Val Leu Asp Pro Ser Asn Pro Leu Thr Trp Pro Leu  
 145 150 155 160  
 Trp Ala Ala Tyr Gly Ala Val Thr Gly Thr Val Ala Met Gly Leu Trp  
 165 170 175  
 Val Leu Ala His Glu Cys Gly His Gly Ala Phe Ser Lys Asn Arg Ser  
 180 185 190  
 Leu Gln Asp Ala Val Gly Tyr Ile Ile His Ser Ile Met Leu Val Pro  
 195 200 205  
 Tyr Phe Ser Trp Gln Arg Ser His Ala Val His His Gln Tyr Thr Asn  
 210 215 220  
 His Met Glu Leu Gly Glu Thr His Val Pro Asp Arg Ala Asp Lys Glu  
 225 230 235 240  
 Gly Glu Lys Ser Leu Ala Leu Arg Gln Phe Met Leu Asp Ser Phe Gly  
 245 250 255  
 Lys Asp Lys Gly Met Lys Ala Tyr Gly Gly Leu Gln Ser Phe Leu His  
 260 265 270  
 Leu Ile Val Gly Trp Pro Ala Tyr Leu Leu Ile Gly Ala Thr Gly Gly  
 275 280 285  
 Pro Asp Arg Gly Met Thr Asn His Phe Tyr Pro Asn Pro Leu Ser Thr  
 290 295 300  
 Pro Thr Gln Pro Lys Lys Glu Leu Phe Pro Gly Asn Trp Lys Glu Lys  
 305 310 315 320  
 Val Tyr Gln Ser Asp Ile Gly Ile Ala Ala Val Val Gly Ala Leu Ile  
 325 330 335  
 Ala Trp Thr Ala Thr Ser Gly Leu Ala Pro Val Met Ala Leu Tyr Gly

340										345										350									
Gly	Pro	Leu	Ile	Val	Ile	Asn	Ala	Trp	Leu	Val	Leu	Tyr	Thr	Trp	Leu														
355										360										365									
Gln	His	Thr	Asp	Thr	Asp	Val	Pro	His	Phe	Ser	Ser	Asp	Asn	His	Asn														
370										375										380									
Phe	Val	Lys	Gly	Ala	Leu	His	Thr	Ile	Asp	Arg	Pro	Tyr	Asp	Lys	Leu														
385										390										395									
Asp	Pro	Trp	Gly	Ile	Ile	Asp	Phe	Leu	His	His	Lys	Ile	Gly	Thr	Thr														
405										410										415									
His	Val	Ala	His	His	Phe	Asp	Ser	Thr	Ile	Pro	His	Tyr	Lys	Ala	Gln														
420										425										430									
Ile	Ala	Thr	Asp	Ala	Ile	Lys	Ala	Lys	Phe	Pro	Glu	Val	Tyr	Leu	Tyr														
435										440										445									
Asp	Pro	Thr	Pro	Ile	Pro	Gln	Ala	Met	Trp	Arg	Val	Ala	Lys	Gly	Cys														
450										455										460									
Thr	Ala	Val	Glu	Gln	Arg	Gly	Asp	Ala	Trp	Val	Trp	Lys	Asn	Glu	Gly														
465										470										475									
Ile	Glu	Asp	Leu	Val	Glu	His	Arg	Gln	Ser	Lys	Leu	Ser	Ser	Glu															
485										490										495									

<210> 7  
<211> 1578  
<212> DNA  
<213> Physcomitrella patens

```
<220>  
<221> CDS  
<222> (1)..(1578)
```

[illegible]

65	70	75	80	
act gcg gag gca ctc gca gaa tca gtc gtg aag ccc acg aga cga agg				288
Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg	85	90	95	
tca tct cag tgg aag aag tgc aca cac ccc cta tca gaa gta gca gta				336
Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val	100	105	110	
cac aac aag cca agc gat tgc tgg att gtt gta aaa aac aag gtg tat				384
His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr	115	120	125	
gat gtt tcc aat ttt gcg gac gag cat ccc gga gga tca gtt att agt				432
Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser	130	135	140	
act tat ttt gga cga gac ggc aca gat gtt ttc tct agt ttt cat gca				480
Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala	145	150	155	160
gct tct aca tgg aaa att ctt caa gac ttt tac att ggt gac gtg gag				528
Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu	165	170	175	
agg gtg gag ccg act cca gag ctg ctg aaa gat ttc cga gaa atg aga				576
Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg	180	185	190	
gct ctt ttc ctg agg gag caa ctt ttc aaa agt tgc aaa ttg tac tat				624
Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr	195	200	205	
gtt atg aag ctg ctc acg aat gtt gct att ttt gct gcg agc att gca				672
Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala	210	215	220	
ata ata tgt tgg agc aag act att tca gcg gtt ttg gct tca gct tgt				720
Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys	225	230	235	240
atg atg gct ctg tgt ttc caa cag tgc gga tgg cta tcc cat gat ttt				768
Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe	245	250	255	
ctc cac aat cag gtg ttt gag aca cgc tgg ctt aat gaa gtt gtc ggg				816
Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly	260	265	270	
tat gtg atc ggc aac gcc gtt ctg ggg ttt agt aca ggg tgg tgg aag				864
Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys	275	280	285	
gag aag cat aac ctt cat cat gct gct cca aat gaa tgc gat cag act				912
Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr	290	295	300	

## 15

tac caa cca att gat gaa gat att gat act ctc ccc ctc att gcc tgg	960
Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp	
305 310 315 320	
agc aag gac ata ctg gcc aca gtt gag aat aag aca ttc ttg cga atc	1008
Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile	
325 330 335	
ctc caa tac cag cat ctg ttc ttc atg ggt ctg tta ttt ttc gcc cgt	1056
Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg	
340 345 350	
ggg agt tgg ctc ttt tgg agc tgg aga tat acc tct aca gca gtg ctc	1104
Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu	
355 360 365	
tca cct gtc gac agg ttg ttg gag aag gga act gtt ctg ttt cac tac	1152
Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr	
370 375 380	
ttt tgg ttc gtc ggg aca gcg tgc tat ctt ctc cct ggt tgg aag cca	1200
Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro	
385 390 395 400	
tta gta tgg atg gcg gtg act gag ctc atg tcc ggc atg ctg ctg ggc	1248
Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly	
405 410 415	
ttt gta ttt gta ctt agc cac aat ggg atg gag gtt tat aat tgc tct	1296
Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser	
420 425 430	
aaa gaa ttc gtg agt gca cag atc gta tcc aca cgg gat atc aaa gga	1344
Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly	
435 440 445	
aac ata ttc aac gac tgg ttc act ggt ggc ctt aac agg caa ata gag	1392
Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu	
450 455 460	
cat cat ctt ttc cca aca atg ccc agg cat aat tta aac aaa ata gca	1440
His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala	
465 470 475 480	
cct aga gtg gag gtg ttc tgt aag aaa cac ggt ctg gtg tac gaa gac	1488
Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp	
485 490 495	
gta tct att gct acc ggc act tgc aag gtt ttg aaa gca ttg aag gaa	1536
Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu	
500 505 510	
gtc gcg gag gct gcg gca gag cag cat gct acc acc agt taa	1578
Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser	
515 520 525	

&lt;211&gt; 525

&lt;212&gt; PRT

&lt;213&gt; Physcomitrella patens

&lt;400&gt; 8

Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
 1 5 10 15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
 20 25 30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
 35 40 45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
 50 55 60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
 65 70 75 80

Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
 85 90 95

Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
 100 105 110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
 115 120 125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
 130 135 140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
 145 150 155 160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
 165 170 175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
 180 185 190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
 195 200 205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
 210 215 220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys  
 225 230 235 240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe  
 245 250 255

Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly  
 260 265 270

Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys  
 275 280 285

Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr  
 290 295 300  
 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp  
 305 310 315 320  
 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile  
 325 330 335  
 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg  
 340 345 350  
 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu  
 355 360 365  
 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr  
 370 375 380  
 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro  
 385 390 395 400  
 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly  
 405 410 415  
 Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser  
 420 425 430  
 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly  
 435 440 445  
 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu  
 450 455 460  
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala  
 465 470 475 480  
 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp  
 485 490 495  
 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu  
 500 505 510  
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser  
 515 520 525

<210> 9  
 <211> 873  
 <212> DNA  
 <213> Physcomitrella patens

<220>  
 <221> CDS  
 <222> (1)..(873)

<400> 9  
 atg gag gtc gtg gag aga ttc tac ggt gag ttg gat ggg aag gtc tcg 48

Met	Glu	Val	Val	Glu	Arg	Phe	Tyr	Gly	Glu	Leu	Asp	Gly	Lys	Val	Ser	
1				5					10					15		
cag	ggc	gtg	aat	gca	ttg	ctg	ggt	agt	ttt	ggg	gtg	gag	ttg	acg	gat	96
Gln	Gly	Val	Asn	Ala	Leu	Leu	Gly	Ser	Phe	Gly	Val	Glu	Leu	Thr	Asp	
			20					25					30			
acg	ccc	act	acc	aaa	ggc	ttg	ccc	ctc	gtt	gac	agt	ccc	aca	ccc	atc	144
Thr	Pro	Thr	Thr	Lys	Gly	Leu	Pro	Leu	Val	Asp	Ser	Pro	Thr	Pro	Ile	
			35				40					45				
gtc	ctc	ggt	gtt	tct	gta	tac	ttg	act	att	gtc	att	gga	ggg	ctt	ttg	192
Val	Leu	Gly	Val	Ser	Val	Tyr	Leu	Thr	Ile	Val	Ile	Gly	Gly	Leu	Leu	
			50				55					60				
tgg	ata	aag	gcc	agg	gat	ctg	aaa	ccg	cgc	gcc	tcg	gag	cca	ttt	ttg	240
Trp	Ile	Lys	Ala	Arg	Asp	Leu	Lys	Pro	Arg	Ala	Ser	Glu	Pro	Phe	Leu	
	65				70					75					80	
ctc	caa	gct	ttg	gtg	ctt	gtg	cac	aac	ctg	ttc	tgt	ttt	gcg	ctc	agt	288
Leu	Gln	Ala	Leu	Val	Leu	Val	His	Asn	Leu	Phe	Cys	Phe	Ala	Leu	Ser	
				85					90					95		
ctg	tat	atg	tgc	gtg	ggc	atc	gct	tat	cag	gct	att	acc	tgg	cgg	tac	336
Leu	Tyr	Met	Cys	Val	Gly	Ile	Ala	Tyr	Gln	Ala	Ile	Thr	Trp	Arg	Tyr	
			100					105					110			
tct	ctc	tgg	ggc	aat	gca	tac	aat	cct	aaa	cat	aaa	gag	atg	gcg	att	384
Ser	Leu	Trp	Gly	Asn	Ala	Tyr	Asn	Pro	Lys	His	Lys	Glu	Met	Ala	Ile	
		115					120					125				
ctg	gta	tac	ttg	ttc	tac	atg	tct	aag	tac	gtg	gaa	ttc	atg	gat	acc	432
Leu	Val	Tyr	Leu	Phe	Tyr	Met	Ser	Lys	Tyr	Val	Glu	Phe	Met	Asp	Thr	
			130				135					140				
gtt	atc	atg	ata	ctg	aag	cgc	agc	acc	agg	caa	ata	agc	ttc	ctc	cac	480
Val	Ile	Met	Ile	Leu	Lys	Arg	Ser	Thr	Arg	Gln	Ile	Ser	Phe	Leu	His	
	145				150					155				160		
gtt	tat	cat	cat	tct	tca	att	tcc	ctc	att	tgg	tgg	gct	att	gct	cat	528
Val	Tyr	His	His	Ser	Ser	Ile	Ser	Leu	Ile	Trp	Trp	Ala	Ile	Ala	His	
				165					170					175		
cac	gct	cct	ggc	ggt	gaa	gca	tat	tgg	tct	gcg	gct	ctg	aac	tca	gga	576
His	Ala	Pro	Gly	Gly	Glu	Ala	Tyr	Trp	Ser	Ala	Ala	Leu	Asn	Ser	Gly	
			180					185					190			
gtg	cat	gtt	ctc	atg	tat	gcg	tat	tac	ttc	ttg	gct	gcc	tgc	ctt	cga	624
Val	His	Val	Leu	Met	Tyr	Ala	Tyr	Tyr	Phe	Leu	Ala	Ala	Cys	Leu	Arg	
			195				200					205				
agt	agc	cca	aag	tta	aaa	aat	aag	tac	ctt	ttt	tgg	ggc	agg	tac	ttg	672
Ser	Ser	Pro	Lys	Leu	Lys	Asn	Lys	Tyr	Leu	Phe	Trp	Gly	Arg	Tyr	Leu	
		210				215					220					
aca	caa	ttc	caa	atg	ttc	cag	ttt	atg	ctg	aac	tta	gtg	cag	gct	tac	720
Thr	Gln	Phe	Gln	Met	Phe	Gln	Phe	Met	Leu	Asn	Leu	Val	Gln	Ala	Tyr	
					230					235					240	

tac gac atg aaa acg aat gcg cca tat cca caa tgg ctg atc aag att 768  
 Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
                   245                  250                  255

ttg ttc tac tac atg atc tcg ttg ctg ttt ctt ttc ggc aat ttt tac 816  
 Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
                   260                  265                  270

gta caa aaa tac atc aaa ccc tct gac gga aag caa aag gga gct aaa 864  
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
                   275                  280                  285

act gag tga 873  
 Thr Glu  
           290

<210> 10

<211> 290

<212> PRT

<213> Physcomitrella patens

<400> 10

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser  
   1                  5                  10                  15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp  
                   20                  25                  30

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
                   35                  40                  45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu  
                   50                  55                  60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu  
                   65                  70                  75                  80

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
                   85                  90                  95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr  
                   100                  105                  110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile  
                   115                  120                  125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr  
                   130                  135                  140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His  
                   145                  150                  155                  160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His  
                   165                  170                  175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly

180	185	190
Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg		
195	200	205
Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu		
210	215	220
Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr		
225	230	235 240
Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile		
245	250	255
Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr		
260	265	270
Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys		
275	280	285
Thr Glu		
290		
<210> 11		
<211> 1526		
<212> DNA		
<213> Phaeodactylum tricornutum		
<220>		
<221> CDS		
<222> (92)..(1402)		
<400> 11		
gcttcggtta gcgtcccata gtttgttaca cttggctgtg aaacgaatac gttcttggtc 60		
tacttactac aacgaagcaa ccaccagcag c atg ggt aag gga ggt caa cga 112		
	Met Gly Lys Gly Gly Gln Arg	
	1 5	
gct gta gct ccc aag agt gcc acc agc tct act ggc agt gct acc ctt 160		
Ala Val Ala Pro Lys Ser Ala Thr Ser Ser Thr Gly Ser Ala Thr Leu		
10 15 20		
agc caa agc aag gaa cag gta tgg act tcg tcg tac aac cct ctg gcg 208		
Ser Gln Ser Lys Glu Gln Val Trp Thr Ser Ser Tyr Asn Pro Leu Ala		
25 30 35		
aag gat tcc ccg gag ctg cca acc aaa ggc caa atc aag gcc gtc att 256		
Lys Asp Ser Pro Glu Leu Pro Thr Lys Gly Gln Ile Lys Ala Val Ile		
40 45 50 55		
ccg aag gaa tgt ttc caa cgc tca gcc ttt tgg tct acc ttc tac ctg 304		
Pro Lys Glu Cys Phe Gln Arg Ser Ala Phe Trp Ser Thr Phe Tyr Leu		
60 65 70		
atg cgc gat ctc gcc atg gct gcc gcc ttt tgc tac gga acc tca cag 352		
Met Arg Asp Leu Ala Met Ala Ala Ala Phe Cys Tyr Gly Thr Ser Gln		

75	80	85	
gtc ctc tcc acc gac ctt ccc caa gac gcc acg ctc att ctg ccc tgg Val Leu Ser Thr Asp Leu Pro Gln Asp Ala Thr Leu Ile Leu Pro Trp 90 95 100			400
gct ctc ggc tgg ggc gtc tac gcc ttt tgg atg gga acc att ctc acc Ala Leu Gly Trp Gly Val Tyr Ala Phe Trp Met Gly Thr Ile Leu Thr 105 110 115			448
ggg cct tgg gta gtt gcg cac gaa tgt gga cac ggc gct tac tcc gac Gly Pro Trp Val Val Ala His Glu Cys Gly His Gly Ala Tyr Ser Asp 120 125 130 135			496
tcc cag acg ttc aat gac gtg gtc ggc ttt atc gtc cac caa gct ttg Ser Gln Thr Phe Asn Asp Val Val Gly Phe Ile Val His Gln Ala Leu 140 145 150			544
ctc gtc ccc tac ttt gcc tgg cag tac acc cac gcg aaa cac cac cgt Leu Val Pro Tyr Phe Ala Trp Gln Tyr Thr His Ala Lys His His Arg 155 160 165			592
cga acc aac cat ctg gtg gac ggc gag tcc cac gtc cct tct acc gcc Arg Thr Asn His Leu Val Asp Gly Glu Ser His Val Pro Ser Thr Ala 170 175 180			640
aag gat aac ggc ctc ggg ccg cac aac gag cga aac tcc ttc tac gcc Lys Asp Asn Gly Leu Gly Pro His Asn Glu Arg Asn Ser Phe Tyr Ala 185 190 195			688
gcg tgg cac gag gcc atg gga gac ggc gcc ttt gcc gtc ttt caa gtc Ala Trp His Glu Ala Met Gly Asp Gly Ala Phe Ala Val Phe Gln Val 200 205 210 215			736
tgg tcg cac ttg ttc gtc ggc tgg cct ctc tac ttg gcc ggt ctg gcc Trp Ser His Leu Phe Val Gly Trp Pro Leu Tyr Leu Ala Gly Leu Ala 220 225 230			784
agt acc gga aag ctt gcg cac gaa ggt tgg tgg ctg gaa gaa cgg aac Ser Thr Gly Lys Leu Ala His Glu Gly Trp Trp Leu Glu Glu Arg Asn 235 240 245			832
gcg att gcg gat cac ttt cga ccc agc tct ccc atg ttc ccc gcc aag Ala Ile Ala Asp His Phe Arg Pro Ser Ser Pro Met Phe Pro Ala Lys 250 255 260			880
atc cgt gcc aag att gcc ctt tcc agc gcg acg gaa ctc gcc gtg ctc Ile Arg Ala Lys Ile Ala Leu Ser Ser Ala Thr Glu Leu Ala Val Leu 265 270 275			928
gct gga ctc ttg tat gtc ggt aca cag gtc gga cac ctt ccc gtc ctg Ala Gly Leu Leu Tyr Val Gly Thr Gln Val Gly His Leu Pro Val Leu 280 285 290 295			976
ctg tgg tac tgg gga ccg tac acc ttt gtc aac gct tgg ctt gta ctc Leu Trp Tyr Trp Gly Pro Tyr Thr Phe Val Asn Ala Trp Leu Val Leu 300 305 310			1024

tac acg tgg ctg cag cat acg gac ccg tcc atc ccg cac tac ggt gaa 1072  
 Tyr Thr Trp Leu Gln His Thr Asp Pro Ser Ile Pro His Tyr Gly Glu  
 315 320 325  
 ggc gag tgg acc tgg gtc aag ggc gcg ctc tct acc att gat cga gac 1120  
 Gly Glu Trp Thr Trp Val Lys Gly Ala Leu Ser Thr Ile Asp Arg Asp  
 330 335 340  
 tac ggc atc ttc gat ttc ttt cac cac acc atc ggt tcc acg cac gtg 1168  
 Tyr Gly Ile Phe Asp Phe Phe His His Thr Ile Gly Ser Thr His Val  
 345 350 355  
 gta cac cat ttg ttc cac gaa atg ccc tgg tac aat gcc ggc att gcc 1216  
 Val His His Leu Phe His Glu Met Pro Trp Tyr Asn Ala Gly Ile Ala  
 360 365 370 375  
 acg caa aag gtc aag gaa ttt ttg gaa ccc cag ggc ttg tac aat tac 1264  
 Thr Gln Lys Val Lys Glu Phe Leu Glu Pro Gln Gly Leu Tyr Asn Tyr  
 380 385 390  
 gat ccg acc ccc tgg tac aag gcc atg tgg cgc att gcc cgg acc tgt 1312  
 Asp Pro Thr Pro Trp Tyr Lys Ala Met Trp Arg Ile Ala Arg Thr Cys  
 395 400 405  
 cac tat gtg gag tca aac gag ggt gtg cag tat ttc aag agt atg gaa 1360  
 His Tyr Val Glu Ser Asn Glu Gly Val Gln Tyr Phe Lys Ser Met Glu  
 410 415 420  
 aac gtg ccg ctg act aag gat gtg cga aac aaa gcc gca tga 1402  
 Asn Val Pro Leu Thr Lys Asp Val Arg Asn Lys Ala Ala  
 425 430 435  
 gaaaaagtgc caccgacgca taattttaca atcctaccaa caagaccaac attatatgggt 1462  
 tttcgcttaa aagatagttt tttctaccat ctgtgtagtc ggcacaaaaa aaaaaaaaaa 1522  
 aaaa 1526

&lt;210&gt; 12

&lt;211&gt; 436

&lt;212&gt; PRT

<213> *Phaeodactylum tricornutum*

&lt;400&gt; 12

Met Gly Lys Gly Gly Gln Arg Ala Val Ala Pro Lys Ser Ala Thr Ser  
 1 5 10 15

Ser Thr Gly Ser Ala Thr Leu Ser Gln Ser Lys Glu Gln Val Trp Thr  
 20 25 30

Ser Ser Tyr Asn Pro Leu Ala Lys Asp Ser Pro Glu Leu Pro Thr Lys  
 35 40 45

Gly Gln Ile Lys Ala Val Ile Pro Lys Glu Cys Phe Gln Arg Ser Ala  
 50 55 60

Phe Trp Ser Thr Phe Tyr Leu Met Arg Asp Leu Ala Met Ala Ala Ala

65	70	75	80
Phe Cys Tyr Gly Thr Ser Gln Val Leu Ser Thr Asp Leu Pro Gln Asp	85	90	95
Ala Thr Leu Ile Leu Pro Trp Ala Leu Gly Trp Gly Val Tyr Ala Phe	100	105	110
Trp Met Gly Thr Ile Leu Thr Gly Pro Trp Val Val Ala His Glu Cys	115	120	125
Gly His Gly Ala Tyr Ser Asp Ser Gln Thr Phe Asn Asp Val Val Gly	130	135	140
Phe Ile Val His Gln Ala Leu Leu Val Pro Tyr Phe Ala Trp Gln Tyr	145	150	155
Thr His Ala Lys His His Arg Arg Thr Asn His Leu Val Asp Gly Glu	165	170	175
Ser His Val Pro Ser Thr Ala Lys Asp Asn Gly Leu Gly Pro His Asn	180	185	190
Glu Arg Asn Ser Phe Tyr Ala Ala Trp His Glu Ala Met Gly Asp Gly	195	200	205
Ala Phe Ala Val Phe Gln Val Trp Ser His Leu Phe Val Gly Trp Pro	210	215	220
Leu Tyr Leu Ala Gly Leu Ala Ser Thr Gly Lys Leu Ala His Glu Gly	225	230	235
Trp Trp Leu Glu Glu Arg Asn Ala Ile Ala Asp His Phe Arg Pro Ser	245	250	255
Ser Pro Met Phe Pro Ala Lys Ile Arg Ala Lys Ile Ala Leu Ser Ser	260	265	270
Ala Thr Glu Leu Ala Val Leu Ala Gly Leu Leu Tyr Val Gly Thr Gln	275	280	285
Val Gly His Leu Pro Val Leu Leu Trp Tyr Trp Gly Pro Tyr Thr Phe	290	295	300
Val Asn Ala Trp Leu Val Leu Tyr Thr Trp Leu Gln His Thr Asp Pro	305	310	315
Ser Ile Pro His Tyr Gly Glu Gly Glu Trp Thr Trp Val Lys Gly Ala	325	330	335
Leu Ser Thr Ile Asp Arg Asp Tyr Gly Ile Phe Asp Phe Phe His His	340	345	350
Thr Ile Gly Ser Thr His Val Val His His Leu Phe His Glu Met Pro	355	360	365
Trp Tyr Asn Ala Gly Ile Ala Thr Gln Lys Val Lys Glu Phe Leu Glu	370	375	380

Pro Gln Gly Leu Tyr Asn Tyr Asp Pro Thr Pro Trp Tyr Lys Ala Met  
385 390 395 400

Trp Arg Ile Ala Arg Thr Cys His Tyr Val Glu Ser Asn Glu Gly Val  
405 410 415

Gln Tyr Phe Lys Ser Met Glu Asn Val Pro Leu Thr Lys Asp Val Arg  
420 425 430

Asn Lys Ala Ala  
435

<210> 13

<211> 3598

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche  
Promotor-Terminator-Expressionskassette in Vektor  
pUC19 dar

<400> 13

tgcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatattgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300  
tacgccagct ggcgaaagg ggatgtgctg caaggcgatt aagtgggta acgccagggt 360  
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agtcctcga 420  
gcaaatttac acattgccac taaacgtcta aaccttgta atttgtttt gttttactat 480  
gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540  
tttatgctaa cgtttgcaa cacttagcaa ttgcaagtt gattaattga ttctaaatta 600  
ttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgaga ttttaattgtt 720  
gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780  
taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggttag taatttttca 840  
agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900  
ttaaaaaat tttggaaatg atttgcattg aagccatgtg taaaaccatg acatccactt 960  
ggaggatgca ataataaga aaactacaaa tttacatgca actagttag catgtagtct 1020

atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
 taattttcttc atagccagcc caccgcggtg ggccggccgc tgcagtctag aaggcctcct 1140  
 gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca attctgttgt 1200  
 gcacgttgta aaaaacctga gcatgtgtag ctacagatcct taccgcggtt ttccggttcat 1260  
 tctaataaat atatcaccgg ttactatcgt atttttatga ataataattct ccgttcaatt 1320  
 tactgattgt ccgtcgacga attcgagctc ggccgcgcaa gcttggcgta atcatggtca 1380  
 tagctgtttc ctgtgtgaaa ttgttatccg ctacacaattc cacacaacat acgagccgga 1440  
 agcataaagt gtaaagcctg gggcgcctaa tgagtgcgtc aactcacatt aattgcgttg 1500  
 cgtcactgc ccgttttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 1560  
 caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac 1620  
 tcgctgcgtc cggctgttgc gctgcggcga gcggtatcag ctactcaca ggccgtaata 1680  
 cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 1740  
 aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct 1800  
 gacgagcatc acaaaaaatc acgtcaagt cagaggtggc gaaacccgac aggactataa 1860  
 agataccagg cgtttccccc tggaagctcc ctctgcgct ctctgttcc gaccctgccg 1920  
 cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgcttcc tcatagctca 1980  
 cgctgtaggc atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa 2040  
 cccccgttc agcccgaccg ctgcgcctta tccggttaact atcgtcttga gtccaacccg 2100  
 gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 2160  
 tatgtaggcg gtgctacaga gttcttgaag tggcggccta actacggcta cactagaagg 2220  
 acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 2280  
 tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgttg caagcagcag 2340  
 attacgcgca gaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac 2400  
 gctcagtgga acgaaaactc acgttaaggg attttggta tgagattatc aaaaaggatc 2460  
 ttacactaga tctttttaa ttaaaaatga agttttaaat caatctaaag tatatatgag 2520  
 taaacttggc ctgacagtta ccaatgctta atcagtagg cacctatctc agcgatctgt 2580  
 ctatttcgtt catccatagt tgctgactc ccgctcgtgt agataactac gatacgggag 2640  
 ggcttaccat ctggccccag tgctgcaatg ataccgcgag acccagctc accggctcca 2700  
 gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tcttgcaact 2760

ttatccgcct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca 2820  
 gttaatagtt tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg 2880  
 tttggtatgg cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc 2940  
 atgtttgtgca aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg 3000  
 gccgcagtgt tatcactcat ggttatggca gcactgcata attctcttac tgtcatgccca 3060  
 tccgtaagat gcttttctgt gactggtgag tactcaacca agtcattctg agaatagttg 3120  
 atgcccgcgac cgagttgctc ttgcccggcg tcaataccggg ataataccgc gccacatagc 3180  
 agaactttaa aagtgtcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc 3240  
 ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca 3300  
 tcttttactt tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa 3360  
 aagggataaa gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat 3420  
 tgaagcattt atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa 3480  
 aataa caaaa taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa 3540  
 accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 3598

<210> 14

<211> 3590

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche  
 Promotor-Terminator-Expressionskassette in Vektor  
 pUC19 dar

<400> 14

tcgcgctgtt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
 attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300  
 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360  
 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga 420  
 gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat 480  
 gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540

tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600  
 tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
 tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttgaga ttttaattggt 720  
 gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780  
 taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaaggtttag taatttttca 840  
 agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900  
 ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960  
 ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttatg catgtagtct 1020  
 atataatgag gattttgcaa tactttcatt catcacact cactaagttt tacacgatta 1080  
 taattttctt atagccagcg gatccgatat cgggcccgct agcgtaacc ctgctttaat 1140  
 gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg 1200  
 taaaaaacct gagcatgtgt agctcagatc cttaccgccc gtttcggttc attctaatga 1260  
 atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt 1320  
 gtccgtcgac gaattcgagc tcggcgcgcc aagcttgccg taatcatggt catagctgtt 1380  
 tcctgtgtga aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa 1440  
 gtgtaaagcc tgggggtgcct aatgagttag ctaactcaca ttaattgcgt tgcgctcact 1500  
 gccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 1560  
 ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg 1620  
 ctccgtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa tacggttatc 1680  
 cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 1740  
 gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca 1800  
 tcacaaaaat cgacgctcaa gtcagagggt gcgaaaccg acaggactat aaagatacca 1860  
 ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgttaccgg 1920  
 atacctgtcc gcctttctcc ctccgggaag cgtggcgctt tctcatagct cacgctgtag 1980  
 gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 2040  
 tcagcccgac cgctgcgct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 2100  
 cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 2160  
 cgggtctaca gagttcttga agtgggtggc taactacgc tacactagaa ggacagtatt 2220  
 tggatatctg gctctgctga agccagttac ctccggaaaa agagttggta gctcttgatc 2280

cggcaaacaa accaccgctg gtagcggtag tttttttggt tgcaagcagc agattacgcg 2340  
 cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acggtcagtg 2400  
 gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 2460  
 gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg 2520  
 gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 2580  
 ttcattcata gttgcctgac tccccgtcgt gtagataact acgatacggg agggccttacc 2640  
 atctggcccc agtgctgcaa tgataccgag agaccacgcg tcaccggctc cagattttatc 2700  
 agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 2760  
 ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 2820  
 tttgcgcaac gttgttgcca ttgctacagg catcgtgggt tcacgctcgt cgtttggtat 2880  
 ggcttcattc agctccggtt cccaacgata aaggcgagtt acatgatccc ccatgttggtg 2940  
 caaaaaagcg gtttagctct tcggtcctcc gatcgttggtc agaagtaagt tggccgcagt 3000  
 gttatcactc atgggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 3060  
 atgcttttct gtgactgggt agtactcaac caagtcattc tgagaatagt gtatgcggcg 3120  
 accgagttgc tcttgcccgg cgtcaatagc ggataatacc gcgccacata gcagaacttt 3180  
 aaaagtgtc atcattggaa aacgtttctt ggggcgaaaa ctctcaagga tcttaccgct 3240  
 gttgagatcc agttcgatgt aacccactcg tgcaccaaac tgatcttcag catcttttac 3300  
 tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 3360  
 aaggcgacac cggaaatggt gaatactcat actcttctt tttcaatatt attgaagcat 3420  
 ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 3480  
 aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat 3540  
 tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 3590

<210> 15

<211> 3584

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche  
 Promotor-Terminator-Expressionskassette in Vektor  
 pUC19 dar

<400> 15

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
 attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300  
 tacgccagct ggcgaaaggg ggatgtgtg caaggcgatt aagttgggta acgccagggt 360  
 tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga 420  
 gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat 480  
 gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540  
 tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600  
 tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
 tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt 720  
 gcaatgctgc atggatggca tatacaccaa acattcaata attccttgagg ataataatgg 780  
 taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaagggttag taatttttca 840  
 agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900  
 ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960  
 ggaggatgca ataatgaaga aaactacaaa tttacatgca actagttagt catgtagtct 1020  
 atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
 taattttctt atagccagca gatctgcgg catcgatccc gggccatggc ctgctttaat 1140  
 gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg 1200  
 taaaaaacct gagcatgtgt agctcagatc cttaccgccc gtttcggttc attctaata 1260  
 atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt 1320  
 gtccgtcgac gagctcggcg cgccaagctt ggcgtaatca tggatcatgc tgtttcctgt 1380  
 gtgaaattgt tatccgtca caattccaca caacatacga gccggaagca taaagtgtaa 1440  
 agcctggggg gcctaataag tgagctaact cacattaatt gcgttgcgct cactgcccgc 1500  
 tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag 1560  
 agggcggttg cgtattgggc gctotccgc ttctcgtc actgactcgc tgcgctcgg 1620  
 cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgg tatccacaga 1680  
 atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 1740  
 taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcacia 1800

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggtt 1860  
 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta cgggatacct 1920  
 gtccgccttt ctcccttcgg gaagcggtggc gctttctcat agctcacgct gtaggtatct 1980  
 cagttcggtg taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040  
 cgaccgctgc gccttatccg gtaactatcg tcttgagtc aacccggtaa gacacgactt 2100  
 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtgc 2160  
 tacagagttc ttgaagtggg ggctaacta cggctacact agaaggacag tatttggtat 2220  
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 2280  
 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 2340  
 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 2400  
 aaactcacgt taagggattt tggctatgag attatcaaaa aggatcttca cctagatcct 2460  
 tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttgggtctga 2520  
 cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 2580  
 catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 2640  
 cccagtgct gcaatgatac cgcgagacc acgctcaccg gctccagatt tatcagcaat 2700  
 aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 2760  
 ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 2820  
 caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880  
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940  
 agcgggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttacc 3000  
 actcatgggt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt 3060  
 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 3120  
 ttgctcttgc cggcgctcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180  
 gctcatcatt ggaaaacgtt cttcgggggc aaaactctca aggatcttac cgctgttgag 3240  
 atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac 3300  
 cagcgtttct ggggtgagcaa aaacaggaag gcaaatgcc gcaaaaaagg gaataagggc 3360  
 gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca 3420  
 gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480  
 ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 3540

gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtc

3584

<210> 16

<211> 4507

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche  
Promotor-Terminator-Expressionskassette in Vektor  
pUC19 dar

<400> 16

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatattgc gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
attcgccatt caggtgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300  
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360  
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cggcgcgccg agctcctcga 420  
gcaaatttac acattgccac taaacgtcta aacccttgta atttgttttt gttttactat 480  
gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540  
tttatgctaa cgtttgcaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600  
tttttgcctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatttc 660  
tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga tttaattgtt 720  
gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780  
taccacacaa gatgtgaggt gcatgaacgt cagtgagaca aaagggttag taatttttca 840  
agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900  
ttaaaaaat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatccactt 960  
ggaggatgca ataatagaaga aaactacaaa ttacatgca actagttatg catgtagtct 1020  
atataatgag gattttgcaa tactttcatt catacacact cactaagttt tacacgatta 1080  
taatttcttc atagccagcc caccgcgggtg ggcggccgcc tgcagtctag aaggcctcct 1140  
gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca attctgttgt 1200  
gcaggttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggt ttgggttcat 1260  
tctaataaat atatcaccgc ttactatcgt atttttatga ataataattt cgttcaatt 1320

tactgattgt cgcgcgagca aatttacaca ttgccactaa acgtctaaac ccttgtaatt 1380  
 tgtttttgtt ttactatgtg tggtatgtat ttgatttgcg ataaattttt atatttggt 1440  
 ctaaatttat aacacctttt atgctaacgt ttgccaacac ttagcaattt gcaagttgat 1500  
 taattgattc taaattattt ttgtcttcta aatacatata ctaatcaact ggaaatgtaa 1560  
 atatttgcta atatttctac tataggagaa ttaaagttag tgaatatggt accacaaggt 1620  
 ttggagattt aattgttgca atgctgcatg gatggcatat acaccaaaca ttcaataatt 1680  
 cttgaggata ataattgtac cacacaagat ttgaggtgca tgaacgtcac gtggacaaaa 1740  
 ggtttagtaa tttttcaaga caacaatggt accacacaca agttttgagg tgcattgcat 1800  
 gatgccctgt ggaaagttaa aaaatatatt ggaaatgatt tgcattggaag ccatgtgtaa 1860  
 aaccatgaca tccacttgga ggatgcaata atgaagaaaa ctacaaattt acatgcaact 1920  
 agttatgcat gtagtctata taatgaggat ttgcaatac ttctattcat acacactcac 1980  
 taagttttac acgattataa tttcttcata gccagcggat ccgatatcgg gcccgctagc 2040  
 gttaaccctg ctttaattgag atatgcgaga cgcctatgat ccgatgatat ttgctttcaa 2100  
 ttctgttgtg caggttgtaa aaaacctgag catgtgtagc tcagatcctt accgccggtt 2160  
 tcggttcatt ctaatgaata tatcacccgt tactatcgta tttttatgaa taatattctc 2220  
 cgttcaattt actgattgtc cgcgcagcaa ttgcagctcg gcgcgccaaag cttggcgtaa 2280  
 tcatggcat agctgtttcc tgtgtgaaat tggtatccgc tcacaattcc acacaacata 2340  
 cgagccggaa gcataaagtg taaagcctgg ggtgcctaata gagtgagcta actcacatta 2400  
 attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgcctgtcca gctgcattaa 2460  
 tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg ggcgctcttc cgcttctcgc 2520  
 ctactgact cgcctgcctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag 2580  
 gcggtaatac gggtatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa 2640  
 ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 2700  
 cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 2760  
 ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc tctgttccg 2820  
 accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 2880  
 catagctcac gctgtaggta tctcagttcg gtgtaggctc ttcgctccaa gctgggctgt 2940  
 gtgcacgaac ccccggttca gcccgaccgc tgcgccttat ccggttaact tcgtcttgag 3000  
 tccaaccgg taagacacga cttatcgcca ctggcagcag ccaactggtaa caggattagc 3060

agagcgcggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 3120  
 actagaagga cagtatttgg tatctgogct ctgctgaagc cagttacctt cggaaaaaga 3180  
 gttggtagct cttgatccgg caaacaacc accgctggta gcggtggtt ttttgtttgc 3240  
 aagcagcaga ttacgcgcag aaaaaagga tctcaagaag atcctttgat cttttctacg 3300  
 gggctctgacg ctcaagtggaa cgaaaactca cgtaagggg ttttggatc gagattatca 3360  
 aaaaggatct tcacctagat ctttttaaataaaaaatgaa gttttaaatc aatctaaagt 3420  
 atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgggc acctatctca 3480  
 gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg 3540  
 atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 3600  
 ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggc 3660  
 cctgcaactt tatccgcctc catccagtct attaattggt gccgggaagc tagagtaagt 3720  
 agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtgggtgtca 3780  
 cgctcgtcgt ttgggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca 3840  
 tgatccccc tgttgtgcaa aaaagcgggt agtccttcg gtctccgat cgttgtcaga 3900  
 agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact 3960  
 gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga 4020  
 gaatagtgtg tgcggcgacc gagttgctct tgcccgcggt caatacggga taataccgcg 4080  
 ccacatagca gaactttaaa agtgcctatc attggaaaac gttcttcggg gcgaaaactc 4140  
 tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga 4200  
 tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat 4260  
 gccgcacaaa aggggaataag ggcgacacgg aaatgttgaa tactcatact ctcccttttt 4320  
 caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt 4380  
 atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac 4440  
 gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat cagcaggccc 4500  
 tttcgtc . 4507

<210> 17

<211> 5410

<212> DNA

<213> Unknown

<220>

<223> Sequenz stellt eine pflanzliche  
Promotor-Terminator-Expressionskassette in Vektor  
pUC19 dar

<400> 17

```

ttttggaaat gatttgcacg gaagccatgt gtaaaacccat gacatccact tggaggatgc 60
aataatgaag aaaactacaa atttacatgc aactagttat gcatgtagtc tatataatga 120
ggattttgca atactttcat tcatacacac tctaagtt ttacacgatt ataatttctt 180
catagccagc ggatccgata tcggggccgc tagcgtaac cctgctttaa tgagatatgc 240
gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc 300
tgagcatgtg tagctcagat ccttaccgcc ggttcgggtt cattctaata aatataatcac 360
ccgttactat cgtattttta tgaataatat tctccgttca atttactgat tgtccgtcga 420
gcaaatttac acattgccac taaacgtcta aacccttgta atttgTTTTT gttttactat 480
gtgtgttatg tatttgattt gcgataaatt tttatatttg gtactaaatt tataacacct 540
tttatgctaa cgtttgccaa cacttagcaa tttgcaagtt gattaattga ttctaaatta 600
tttttgtctt ctaaatacat atactaatca actggaaatg taaatatttg ctaatatctc 660
tactatagga gaattaaagt gagtgaatat ggtaccacaa ggtttggaga ttaattggtt 720
gcaatgctgc atggatggca tatacaccaa acattcaata attcttgagg ataataatgg 780
taccacacaa gatttgaggt gcatgaacgt cacgtggaca aaagggttag taatttttca 840
agacaacaat gttaccacac acaagttttg aggtgcatgc atggatgcc tgtggaaagt 900
ttaaaaatat tttggaaatg atttgcatgg aagccatgtg taaaaccatg acatocactt 960
ggaggatgca ataataaga aaactacaaa tttacatgca actagttatg catgtagtct 1020
atataatgag gattttgcaa tactttcatt catcacact cactaagttt tacacgatta 1080
taatttcttc atagccagca gatctgccg catcgatccc gggccatggc ctgctttaat 1140
gagatatgcg agacgcctat gatcgcatga tatttgcttt caattctgtt gtgcacgttg 1200
taaaaaacct gagcatgtgt agctcagatc cttaccgccg gtttcgggtc attctaata 1260
atatatcacc cgttactatc gtatttttat gaataatatt ctccgttcaa tttactgatt 1320
gtccgtcgac gagctcggcg cgccaagctt ggcgtaata tggatcatagc tgtttcctgt 1380
gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa 1440
agcctggggt gcctaataag tgagctaact cacattaatt gcgttgcgct cactgcccgc 1500
tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag 1560
aggcggtttg cgtattgggc gctcttcgc ttctcgcgc actgactcgc tgcgctcggg 1620

```

cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 1680  
 atcaggggat aacgcaggaa agaacaatgtg agcaaaaaggc cagcaaaagg ccaggaaccg 1740  
 taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia 1800  
 aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt 1860  
 tccccctgga agctccctcg tgcgctctcc tgttcggacc ctgccgctta ccggatacct 1920  
 gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 1980  
 cagttcgggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 2040  
 cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt 2100  
 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtagt taggcgggtg 2160  
 tacagagttc ttgaagtggg gccctaacta cggctacact agaaggacag tatttggtat 2220  
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 2280  
 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 2340  
 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 2400  
 aaactcacgt taagggattt tggcatgag attatcaaaa aggatcttca cctagatcct 2460  
 tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa cttgggtctga 2520  
 cagttaccaaa tgcctaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 2580  
 catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 2640  
 cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat 2700  
 aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 2760  
 ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg 2820  
 caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 2880  
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 2940  
 agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttate 3000  
 actcatgggt atggcagcac tgcataatc tcttactgtc atgccatccg taagatgctt 3060  
 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgatgc ggcgaccgag 3120  
 ttgctcttgc ccggcgtaaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 3180  
 gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag 3240  
 atccagttcg atgtaacca ctctgcacc caactgatct tcagcatctt ttactttcac 3300  
 cagcgtttct ggggtgagcaa aaacaggaag gcaaatgcc gcaaaaaagg gaataagggc 3360

gacacggaaa tgttgaatac tcatactcct cctttttcaa tattattgaa gcatttatca 3420  
 gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 3480  
 ggttcgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 3540  
 gacattaacc tataaaaata ggcgatcac gagggccttt cgtctcgcgc gtttcgggtga 3600  
 tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc 3660  
 ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtggtggcg ggtgtcgggg 3720  
 ctggcttaac tatgcggcat cagagcagat tgtactgaga gtgcaccata tgccgtgtga 3780  
 aataccgcac agatgcgtaa ggagaaaata ccgcatcagg cgccattcgc cattcaggct 3840  
 gcgcaactgt tgggaagggc gatcgggtgc ggctctctcg ctattacgcc agctggcgaa 3900  
 agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg 3960  
 ttgtaaaacg acggccagtg aattcggcgc gccgagctcc tcgagcaaatt ttacacattg 4020  
 ccactaaacg tctaaaccct tgtaatttgt tttgtttta ctatgtgtgt tatgtatttg 4080  
 atttgcgata aatttttata tttggtacta aatttataac accttttatg ctaacgtttg 4140  
 ccaacactta gcaatttgca agttgattaa ttgattctaa attatttttg tcttctaaat 4200  
 acatatacta atcaactgga aatgtaaata tttgctaata tttctactat aggagaatta 4260  
 aagtgagtga atatggtacc acaaggtttg gagatttaatt tgttgcaatg ctgcatggat 4320  
 ggcatacata ccaaaccattc aataattctt gaggataata atggtaaccac acaagatttg 4380  
 aggtgcatga acgtcacgtg gacaaaaggt ttagtaattt ttcaagacaa caatgttacc 4440  
 acacacaagt tttgaggtgc atgcatggat gccctgtgga aagtttataa atattttgga 4500  
 aatgatttgc atggaagcca tgtgtaaaac catgacatcc acttgaggga tgcaataatg 4560  
 aagaaaacta caaatttaca tgcaactagt tatgcatgta gtctatataa tgaggatttt 4620  
 gcaatacttt cattcataca cactcactaa gttttacacg attataattt cttcatagcc 4680  
 agcccaccgc ggtgggcggc cgcctgcagt ctagaaggcc tcctgcttta atgagatatg 4740  
 cgagacgcct atgatcgcat gatatttgc tccaattctg ttgtgcacgt tgtaaaaaac 4800  
 ctgagcatgt gtagctcaga tccttaccgc cggtttcggg tcattctaatt gaatatatca 4860  
 cccgttacta tcgtattttt atgaataata ttctccgttc aatttactga ttgtccgtcg 4920  
 agcaaattta cacattgcca ctaaactgtc aaacccttgt aatttgtttt tggtttacta 4980  
 tgtgtgttat gtatttgatt tgcgataaat ttttatattt ggtactaaat ttataacacc 5040  
 ttttatgcta acgtttgcca acacttagca atttgcaagt tgattaattg attctaaatt 5100

atttttgtct tctaaatata tataactaatc aactggaaat gtaaataattt gctaataattt 5160  
 ctactatagg agaattaaag tgagtgaata tgggtaccaca aggtttggag atttaattgt 5220  
 tgcaatgctg catggatggc atatacacca aacattcaat aattcttgag gataataatg 5280  
 gtaccacaca agatttgagg tgcatagaacg tcacgtggac aaaaggttta gtaatttttc 5340  
 aagacaacaa tgttaccaca cacaagtttt gaggtgcatg catggatgcc ctgtggaaag 5400  
 tttaaaaata 5410

<210> 18  
 <211> 648  
 <212> DNA  
 <213> *Phaeodactylum tricornutum*

<220>  
 <221> CDS  
 <222> (1)..(648)

<220>  
 <223>

<400> 18  
 tgg tgg aaa aac aag cac aac gga cac cac gcc gtc ccc aac ctc cac 48  
 Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His  
 1 5 10 15  
 tgc tcc tcc gca gtc gcg caa gat ggg gac ccg gac atc gat acc atg 96  
 Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met  
 20 25 30  
 ccc ctt ctc gcc tgg tcc gtc cag caa gcc cag tct tac cgg gaa ctc 144  
 Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu  
 35 40 45  
 caa gcc gac gga aag gat tcg ggt ttg gtc aag ttc atg atc cgt aac 192  
 Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn  
 50 55 60  
 caa tcc tac ttt tac ttt ccc atc ttg ttg ctc gcc cgc ctg tcg tgg 240  
 Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp  
 65 70 75 80  
 ttg aac gag tcc ttc aag tgc gcc ttt ggg ctt gga gct gcg tcg gag 288  
 Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu  
 85 90 95  
 aac gct gct ctc gaa ctc aag gcc aag ggt ctt cag tac ccc ctt ttg 336  
 Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu  
 100 105 110  
 gaa aag gct ggc atc ctg ctg cac tac gct tgg atg ctt aca gtt tcg 384  
 Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser  
 115 120 125

tcc ggc ttt gga cgc ttc tcg ttc gcg tac acc gca ttt tac ttt cta 432  
 Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu  
 130 135 140

acc gcg acc gcg tcc tgt gga ttc ttg ctc gcc att gtc ttt ggc ctc 480  
 Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu  
 145 150 155 160

ggc cac aac ggc atg gcc acc tac aat gcc gac gcc cgt ccg gac ttc 528  
 Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe  
 165 170 175

tgg aag ctc caa gtc acc acg act cgc aac gtc acg ggc gga cac ggt 576  
 Trp Lys Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly  
 180 185 190

ttc ccc caa gcc ttt gtc gac tgg ttc tgt ggt ggc ctc cag tac caa 624  
 Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln  
 195 200 205

gtc gac cac cac tta ttc ccc agc 648  
 Val Asp His His Leu Phe Pro Ser  
 210 215

<210> 19

<211> 216

<212> PRT

<213> *Phaeodactylum tricornutum*

<400> 19

Trp Trp Lys Asn Lys His Asn Gly His His Ala Val Pro Asn Leu His  
 1 5 10 15

Cys Ser Ser Ala Val Ala Gln Asp Gly Asp Pro Asp Ile Asp Thr Met  
 20 25 30

Pro Leu Leu Ala Trp Ser Val Gln Gln Ala Gln Ser Tyr Arg Glu Leu  
 35 40 45

Gln Ala Asp Gly Lys Asp Ser Gly Leu Val Lys Phe Met Ile Arg Asn  
 50 55 60

Gln Ser Tyr Phe Tyr Phe Pro Ile Leu Leu Leu Ala Arg Leu Ser Trp  
 65 70 75 80

Leu Asn Glu Ser Phe Lys Cys Ala Phe Gly Leu Gly Ala Ala Ser Glu  
 85 90 95

Asn Ala Ala Leu Glu Leu Lys Ala Lys Gly Leu Gln Tyr Pro Leu Leu  
 100 105 110

Glu Lys Ala Gly Ile Leu Leu His Tyr Ala Trp Met Leu Thr Val Ser  
 115 120 125

Ser Gly Phe Gly Arg Phe Ser Phe Ala Tyr Thr Ala Phe Tyr Phe Leu  
 130 135 140

Thr Ala Thr Ala Ser Cys Gly Phe Leu Leu Ala Ile Val Phe Gly Leu  
145 150 155 160

Gly His Asn Gly Met Ala Thr Tyr Asn Ala Asp Ala Arg Pro Asp Phe  
165 170 175

Trp Lys Leu Gln Val Thr Thr Thr Arg Asn Val Thr Gly Gly His Gly  
180 185 190

Phe Pro Gln Ala Phe Val Asp Trp Phe Cys Gly Gly Leu Gln Tyr Gln  
195 200 205

Val Asp His His Leu Phe Pro Ser  
210 215

<210> 20

<211> 12093

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit einer  
Promotor-Terminator-Expressionskassette

<400> 20

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60

gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcc 120

tagtgggagg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180

ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240

atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300

ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360

tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420

gcgtagacgg tctgacgaca cgaaaactgg cggaacgggt gggggttcag cagccggcgc 480

tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg 540

cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600

ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660

ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720

gcgaggcggg tttttcggcc ggggacgcgc tcaatgcgct gatgacaatc agctacttca 780

ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgcca 840

ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttcgacgaag 900

ccggtccgga cgcagcgctc gagcagggac tcgcggtgat tgctgatgga ttggcgaaaa 960

ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgcgtc agacgcccggt agcagcccg c tacgggcttt ttcattgccct gccctagcgt 1140  
 ccaagcctca cggcgcgcgt cggcctctct ggccgccttc tggcgctctt ccgcttcttc 1200  
 gctcactgac tcgctgcgct cggctcgttc gctgcggcga gcggtatcag ctactcaaa 1260  
 ggccgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggcagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagagggtggc gaaaccgcac 1440  
 aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgt ctctgttcc 1500  
 gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata acctgtctc ggggtcatta tagcgatttt ttccgtatat ccatcctttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttcttgg tgtatccaa 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccaccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 aggcggcgcc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggttaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggccgcct ggccggcctg ctgaaactct ggctaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc cactatctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggcat gatggcggtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaag cagctccca tgcgctccat caagaagagc 2280  
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctggttg ccctgcgcg tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgcgctcga agcgtgtgc gagacaccgc ggccggcggc gttgtggata 2460  
 cctcgcggaa aacttgccc tctactgacag atgagggcg gacgttgaca cttgaggggc 2520  
 cgactaccc ggccggcggt tgacagatga gggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700

tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggtt 2820  
 ccgcccgttt ttccggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 aaaccttggt tttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctcgaaccc tcccggcccc ctaacgcggg cctcccatcc cccagggggc 3000  
 tgcgccccct ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 ggggcctggg ttggggcctg cctttcactt cggcgctcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcagggtt tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca ggggcagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc cccgacatag cccactggt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgcccc taatgcgggc tgttgcccg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440

tggtttcaaa atcggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctgggtatt taagggttta gaatgcaagg aacagtgaat tggagtctgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggc cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcctc 4980  
 aggtcttttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccggc gagctgtatg attttttaa gacggaaaag 5160  
 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtgtat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
 ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgTTTTtagta ctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcatca agtgttttg ctctcaggcc gagggccacg gcaagtattt 5520  
 gggcaagggg tcgctgggat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaac aggaataagg gcacattgcc ccggcgtgag tccgggcaat 5700  
 cccgcaagga ggggtgaatga atcggaagtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttcagc ccgtcggctc gatgggtccg caagctacgg ccaagatcga 5880  
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccacggccg ccgtggagcg 5940  
 ttgcgctcgt ctgcaacagg aggcggcagg tttggcgaa tcgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaataac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttcttgttc gatattgcgc cgtggccgga cagatgccga gcgatgcaa acgacacggc 6180

cegctctgcc ctgttcacca cgcgcaacaa gaaaatcccc cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctcg agctgcgggc 6300  
 cgacgatgac gaactgggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aaccttccgc ctcattgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct ggggtcaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtgggtgca gttccggtg ggggttcagc 6900  
 agccagcgtt ttactggcat ttcaggaaca agcgggcact gtcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgtc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattoga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc 7200  
 ggcgcctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaac gcgaggccga 7320  
 ggggtcgccg gtatgtctgt gcgggcgttg cggcggggtt tattgtctgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcatttca tcctcggcgc acttaatt 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcgcg 7500  
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcattc ctgccgtctc gctaggtagc 7560  
 ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggccggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctccgt gcctctgctc 7740  
 acctttaccg cctggcaact ggcgccgga ggacttctgc tcgttccagt agctttagt 7800  
 tttgatccgc caatccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tgggtccggg ggatctcgcg actcgaacct 7920

acagttgttt ccttactggg ctttctcagc cccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcgggtg agcaatggat 8040  
 aggggagttg atatcgtaaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgtgcagc ccgtcccgga ctgatgggct 8400  
 gcctgtatcg agtgggtgatt ttgtgccgag ctgccggtcg gggagctggt ggctggctgg 8460  
 tggcaggata tattgtgggtg taaacaaatt gacgcttaga caacttaata acacattgag 8520  
 gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgcttgccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atgggtggtc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggg tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gttttttggg gtgcaggtgc cgtaaagcac taaatcgga 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaaac tggcgagaaa 8940  
 ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtgcg ggccctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataatt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgcatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660

agttcggctg ggcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgtcogatg cgatgtttcg cttggtgggc gaatgggcag 9780  
 gtagccgat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 tcccttcccc cttcagtgc aacgtcgagc acagctggc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcacccga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctcgctg gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatacgcc aatagcctct ccaccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgacctagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacggt gcggttctgt cagtccaaa cgtaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacatgac ttgatccct 10500  
 ggcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttccaac 10560  
 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgctt 10680  
 ttcccttctc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctcgcg 10740  
 actggcttct tacgtgttcc gcttccctta gcagcccttg cgccttgagt gcttgccgca 10800  
 gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980  
 tttgccaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtctct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgc aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgcctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400

aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagcattata atttcttcat 11520  
 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580  
 tatgcgagac gcctatgac gcctgatatt tgctttcaat tctgttggtc acgttgtaaa 11640  
 aaacctgagc atgtgtagct cagatcctta ccgcggttt cgggttcattc taatgaatat 11700  
 atcaccggtt actatogtat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760  
 gtgcgagaa tcgagctcgg ccgcctcta gaggatcgat gaattcagat cggctgagtg 11820  
 gctccttcaa cgttgccggtt ctgtcagttc caaacgtaaa acggcttggtc ccgcgtcctc 11880  
 ggccgggggtc ataacgtgac tcccttaatt ctccgctcat gatcagattg tcgtttcccg 11940  
 ccttcagttt aaactatcag tgtttgacag gatataattg cgggtaaacc taagagaaaa 12000  
 gagcggttat tagaataatc ggatatttaa aaggcggtga aaagggttat ccttcgtcca 12060  
 tttgtatgtg catgccaacc acagggttcc cca 12093

<210> 21

<211> 12085

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit einer  
Promotor-Terminator-Expressionskassette.

<400> 21

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcc 120  
 tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
 ataatcaggc cgatgcgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatcc gtgcccctt ggacctgttg aacgaggtcg 420  
 gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc 480  
 tttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgcccagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660

ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttctct 720  
 gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gaggagcccg gcgacagcga tgccggcgag cgcggcgcca 840  
 ccgttgaaca ggctccgctc tcgccgctgt tgcggggccg gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcggtc gaggagggac tcgcggtgat tgcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgcgtc agacgcccg agcagcccg tcgggctttt ttcatgccct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttctct 1200  
 gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctactcaaa 1260  
 ggcggtaata cggttatcca cagaatcagg ggataacga ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggcgcgcttg ctggcgcttt tccataggct 1380  
 ccgccccctc gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440  
 aggactataa agataccagg cgtttcccc tggaagctcc ctctgctgct ctctgttcc 1500  
 gacctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatactttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttccttgg tgtatccaa 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttccacctg gcggtgctca acgggaatcc tgctctgca ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggcgccgc ggcggcgatg agcctgtcgg cctacctgct ggccgtcggc cagggcgaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggcccctt gggcgccctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc caccatctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggatcat gatggcgctg gtccgccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgcyg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400

cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata 2460  
 cctcgcgga aacttggccc tctactgacag atgaggggag gacgttgaca cttgaggggc 2520  
 cgactcaccg ggccggcggt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgagagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700  
 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgcccgttt ttccggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 aaaccttggt ttttaaccagg gctgcgcctt gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctcgaaccc tcccggcccg ctaacgcggg cctcccatcc cccagggggc 3000  
 tgcccccctc ggccgcgaac ggctcaccg caaaaatggc agcgctggca gtccttgcca 3060  
 ttgcccggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcacgcacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 gcggcctggg tggcggcctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcatte ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcagggtt tgccgctcaa ttccgtcgt atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgcgc aacaaccgtc ttccggagac tgtcatacgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140

cgtgttgagg ccaacgcccc taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200  
 catatcaatg attttcttggg gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt cgggcgggtgc ttttgccgtt 4320  
 acgcaccacc cgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 tggtttcaaa atcggctccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctggt ttctgggtatt taagggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgtttataa ttagcttctt ggggtatctt taaatactgt agaāaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcacgggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggatatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtggag ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtcttttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
 ccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
 ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcacatc agtgttttgg ctctcaggcc gaggccacg gcaagtattt 5520  
 gggcaagggg tcgctgggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga ggggtgaatga atcggacggt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttcagtc cgtcggctc gatgggtccag caagctacgg ccaagatcga 5880

gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940  
 ttcgcgtcgt ctgcaacagg aggcggcagg tttggcgaag tcgatgacca togacacgcg 6000  
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttccttgctc gatattgcgc cgtggccgga cacgatgcca gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaaca gaaaatcccg cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 cgacgatgac gaactgggtg ggcagcaggt gttggagtac gcaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt ccggttgcca ggtcctgac gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgcgcac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgcg tcaagctgga 6720  
 aaccttccgc ctcatgtgcg gatcggattc caccgcgctg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct gggatcaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtgggtgca gttccggctg ggggttcagc 6900  
 agccagcgtt ttactggcat ttcaggaaca agcgggcact gtcgacgca cttgcttcgc 6960  
 tcagtatcgc tggggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tggtcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggg tgcgagatgc cgtggcattc 7200  
 ggcgcctaca tcgacggcga gatcattggg ctgtcgggtc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaac gcgaggccga 7320  
 ggggtcgcgc gtatgtcgtc gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcatttca tctcggcgc acttaattatt 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggtc gtgttcattc ctgccgtct gctaggtagc 7560  
 ccgatacgat tgatggcggc cctgggggct atttgcgga ctgcgggcgt ggcgctgttg 7620

gtgttgacac caaacgcagc gctagatcct gtcggcgctcg cagcgggcct ggcgggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccggt gcctctgctc 7740  
 acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttctct tggttccggg ggatctcgcg actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc cccagatctg gggctcgatca gccgggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagtgc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct 8400  
 gcctgtatcg agtgggtgatt ttgtgccgag ctgccggctg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtgggtg taaacaaatt gagccttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtgggttttc ttttaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atgggtgggtc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcgga 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcggtgcg ggctctctcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaagc acggccagtg 9120  
 aattaattcc catcttgaag gaaatatagt ttaaataatt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360

agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagtct 9480  
 tcagcaatat cacgggtagc caacgctatg tctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaa 9660  
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctgctcctgc agttcattca gggcacccga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgccc gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aacccgcggc 10380  
 tgagtggctc cttcaacggt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcctgac ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcccccag ctggcaattc cggttcgctt gctgtccata aaaccgccc 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgttgcgtt 10680  
 ttcccttgc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgagg 10740  
 actggcttct tacgtgttcc gcttccttta gcagcccttg cgcctgagt gcttgccgca 10800  
 gcgtgaagct tgcatgcctg caggtcgagc gcgcgccgag ctccctgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980  
 tttgccaaca cttagcaatt tgcaagtga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100

attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgaggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca ttagtctat ataagagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cacgattata atttcttcat 11520  
 agccagcgga tccgatatcg ggcccgctag cgttaaccct gctttaatga gatagcgag 11580  
 acgcctatga tgcgatgata tttgctttca attctgttgt gcacgttgta aaaaacctga 11640  
 gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaatgaat atatcaccgg 11700  
 ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt ccgtcgacga 11760  
 attcgagctc ggcgcgcctc tagaggatcg atgaattcag atcggtgag tggctccttc 11820  
 aacgttgagg ttctgtcagt tccaaacgta aaacggcctt tcccgcgta tcggcgggg 11880  
 tcataacgtg actcccttaa ttctccgctc atgatcagat tgcggtttcc cgccttcagt 11940  
 ttaaaactatc agtgtttgac aggatattt ggccgggtaaa cctaagagaa aagagcggtt 12000  
 attagaataa tcggatattt aaaagggcgt gaaaagggtt atccttcgtc catttgtatg 12060  
 tgcattgcaa ccacagggtt ccca 12085

<210> 22  
 <211> 12079  
 <212> DNA  
 <213> Unknown

<220>  
 <223> pflanzlicher Expressionsvektor mit einer  
 Promotor-Terminator-Expressionskassette

<400> 22  
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 gcgcccagca caggtgcgca ggcaaattgc accaaccgat acagcgccag cagaatgcca 120  
 tagtggggcg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
 ataatacaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaaagt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420

gcgtagacgg tctgacgaca cgcaaactgg cggaacgggtt ggggggttcag cagccggcgc 480  
 ttacttgga cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttctct 720  
 gcgaggcggg tttttcggcc ggggacgcgc tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga 840  
 ccgttgaaca ggctccgctc tcgcgcgtgt tgccggccgc gatagacgcc ttgcacgaag 900  
 ccggtccgga cgcagcgctt gagcagggac tcgcggtgat tgcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgcgtc agacgcccg agcagccgcg tacgggcttt ttcatgccct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggccgcttc tggcgctctt ccgcttctc 1200  
 gctcactgac tcgctgcgct cggctcgttc gctgcggcga gcggtatcag ctactcaaa 1260  
 ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagagggtgc gaaacccgac 1440  
 aggactataa agataccagg cgtttcccc tggaagctcc ctctgcgct ctctgttcc 1500  
 gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttgggtatat ccatectttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttcttgg tgtatccaa 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggcgccgc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggttaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggccgcctc aatggcgacc 2040  
 tgggcgcctt ggccggcctg ctgaaactct ggctaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160

gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgtcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgcgga aacttgccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 cgactcacc ggccgggctg tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgatatt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcggtattga cacttgaggg gcgcgactac 2700  
 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggtt 2820  
 ccgcccgttt ttccggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 aaaccttggt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctogaacco tcccggcccg ctaacgcggg cctcccatcc cccaggggc 3000  
 tgccccctc ggccgcgaac ggctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcagggtgctg gcacgcacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 gcggcctggg tggcggcctg cccttcactt cggccgtcgg ggcattcacg gacttcattg 3240  
 cggggccggc aatttttacc ttgggcatto ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgagggt ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga ctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat ttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcagggtta tgccgctcaa ttcgctgcgt atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900

cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgctc aacaaccgtc ttccggagac tgcatacgc gtaaacacgc cagcgtggc 4020  
 gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgcca taatgcgggc tgttgcccg catccaacgc cattcatggc 4200  
 catatcaatg attttctggg gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc cgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg 4440  
 tggtttcaaa atcggctccg tcgatactat gttatacgc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctgggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggatatata agctggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaagg cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtcttttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggcgg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaa gacggaaaag 5160  
 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340  
 ctatTTTTTg acttaactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgTTTTtagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcatca agtgttttgg ctctcaggcc gaggcccacg gcaagtattt 5520  
 gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640

ggcaccaggc ggggtcaaadc aggaataagg gcacattgcc ccggcgtgag tgggggcaat 5700  
 cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgctgctc 5820  
 gccccgcgaa accttcacgt ccgtcggctc gatggtccag caagctacgg ccaagatcga 5880  
 ggcgcacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940  
 ttgcgctcgt ctcgaaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaanaac cgcggcgag gacctggcaa aacaggctag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttccttggtc gatattgcgc cgtggccgga cagatgcga gcatgcca acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cactgcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 cgacgatgac gaactgggtg ggcagcagg gtgggagtag gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cactccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgctcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgctgccgac 6660  
 ggcccgacgg atgttcgact attcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aaccttcgc ctcatgtgc gatcggattc caccgcgtg aagaagtggc gcgagcagg 6780  
 cggcgaagcc tgcaagagt tgcgaggcag cggcctgggtg gaacacgcct ggggtcaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc 6900  
 agccagcgtt ttactggcat ttcaggaaaca agcgggcact gctcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cactaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc 7200  
 gggcctaca tcgacggcga gatcattggg ctgtcgtct tcaaacagga ggacggcccc 7260  
 aaggacgtc acaaggcgca tctgtccggc gtttctgtg agccgaaca gcgaggccga 7320  
 ggggtcgcgc gtatgctgct gcggcggtt ccggcggtt tattgctcgt gatgatcgtc 7380

cgacagattc caacgggaat ctggtggatg cgcattctca tctcggcgc acttaatt 7440  
 tccgtattct ggagcttggt gtttatttog gtctaccgcc tgcggggcgg ggtcggcggc 7500  
 accgtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc 7560  
 ccgatacgat tgatggcggg cctgggggct atttgcgaa ctgcgggcgt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtgcggctcg cagcgggcct ggcggggcgg 7680  
 gtttccatgg cgttcgggaac cgtgctgacc cgcaagtggc aacctcccg gcctctgctc 7740  
 acctttaccg cctggcaact ggcgccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 tttgatccgc caatccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttctt tggttccggg ggatctcgc actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtccgaa cttcgggtcc ccgacctga ccattcgggt agcaatggat 8040  
 aggggagttg atatcgtaa cgttcacttc taaagaaata gcgccactca gtttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatca cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccg gcagcttagt tgcggttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaa 8700  
 agaatagccc gagatagggt tgagtgttgt tccagtttg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaaac tggcgagaaa 8940  
 ggaagggaag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcggtgcg ggctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagt 9120

aattaattcc catcttgaaa gaaatatagt ttaaattttt attgataaaa taacaagtca 9180  
ggatttatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcatatta 9300  
tgtgtaatac ataaattgat gatatatgta gcttagctca tcgggggatc cgtcgaagct 9360  
agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
tcagcaatat cacgggtagc caacgctatg tcttgatagc ggtccgccac acccagccgg 9540  
ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaa 9660  
agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatgggcag 9780  
gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
agccacgata gccgcgctgc ctgctcctgc agttcattca gggcaccgga caggctggtc 10020  
ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
ccgattgtct gttgtgcccc gtcatagccc aatagcctct ccaccaagc ggccggagaa 10140  
cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcgggc 10380  
tgagtggctc cttaacggtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440  
gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gctcatgatc ttgatccct 10500  
gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcccc 10620  
gtctagctat cgccatgtaa gccactgca agctacctgc tttctcttg cgcttgctt 10680  
ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
actggcttcc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcgga 10800  
gcgtgaagct tgcattgcctg caggctgagc gcgcgccgag ctccctcgagc aaatttacac 10860

attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacy 10980  
 tttgccaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagcattata atttcttcat 11520  
 agccagcaga tctgccggca tcgatcccg gcatggcct gctttaatga gatatgcgag 11580  
 acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga 11640  
 gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaataat atatcaccg 11700  
 ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt ccgtcgacga 11760  
 gctcggcgcg cctctagagg atcgatgaat tcagatcggc tgagtggctc cttcaacgtt 11820  
 gcggttctgt cagttccaaa cgtaaaacgg cttgtcccg gtcacggcg ggggtcataa 11880  
 cgtgactccc ttaattctcc gtcatgac agattgtcgt ttcccgccct cagtttaaac 11940  
 tatcagtgtt tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga 12000  
 ataatcggat atttaaaagg gcgtgaaaag gtttatcctt cgtccatttg tatgtgcatg 12060  
 ccaaccacag ggttcccca 12079

<210> 23

<211> 13002

<212> DNA

<213> Unknown

<220>

<223> pflanzlicher Expressionsvektor mit zwei  
Promotor-Terminator-Expressionskassetten

<400> 23

gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60

gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcc 120

tagtgggacg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacggcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420  
 gcgtagacgg tctgacgaca cgcaaactgg cggaacgggt gggggttcag cagccggcgc 480  
 ttactggca cttcaggaac aagcggggcg tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720  
 gcgaggcggg tttttcggcc ggggacggcg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga 840  
 ccgttgaaca ggctccgctc tcgccgctgt tcggggccgc gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcgctc gagcaggggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgctc agacgcccg agcagccgc tacgggcttt ttcattgcct gccctagcgt 1140  
 ccaagcctca cggccgctc cggcctctct ggccgcttc tggcgctctt ccgttcctc 1200  
 gctcactgac tcgctgcgct cggctcgttc gctcgggcga gcggtatcag ctactcaaa 1260  
 ggcggtataa cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct 1380  
 ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagagggtgc gaaaccgcac 1440  
 aggactataa agataaccagg cgtttcccc tggaagctcc ctctgctcgt ctctgttcc 1500  
 gaccctgccg cttaccggat acctgtccgc cttctccct tcgggaagcg tggcgctttt 1560  
 ccgtgcata accctgcttc ggggtcatta tagcgatttt ttccgtatat ccactcttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttctgttaga ctttccttgg tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggcgccggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgcg acggcgcggt 2100  
 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcggt gattgccaaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgcgga aacttggccc tcaactgacag atgagggggc gacgttgaca cttgaggggc 2520  
 cgactcacc ggcgcggggt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgccct ggggtattga cacttgaggg gcgcgactac 2700  
 tgacagatga ggggcgcgat ccttgacact tgagggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgcccgttt ttcggccacc gctaacctgt cttttaacct gcttttaaac caatatttat 2880  
 aaaccttggt tttaccagg gtgcgccct gtgcgctga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctcgaaacc tcccgcccg ctaacgggg cctcccatcc ccccgagggc 3000  
 tgccccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgcccggat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcacgacat tcagcgacca ggtgcggggc agtgagggcg 3180  
 ggggcctggg tggcggcctg cccttcactt cggcgtcggg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcggtt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cttttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgaaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600

tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat ttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcagggtta tgccgctcaa ttcgctgctt atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcatata gcggccagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgcgc aacaaccgtc ttccggagac tgcatacgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc ccgacatag cccactgtt cgtccatttc cgcgagacg atgacgtcac 4080  
 tgcccggtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgcca taatgcgggc tgttgccgg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccattgttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg 4440  
 tgggtttcaa atcggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggtatata agctgggtgg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggc cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtcttttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag 5160  
 ccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcagggcgga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggaagaaca gtatgtcgag 5340

ctatTTTTtg acttactggg gatcaagcct gattggggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgTTTTtagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcatca agtgTTTTtg ctctcaggcc gagggccacg gcaagtattt 5520  
 gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc gggTcaaTc aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga gggTgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgTgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatggTccag caagctacgg ccaagatcga 5880  
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcgggcg ccgtggagcg 5940  
 ttcgcgTcgt ctgaaacagg aggcggcagg tttggcgaaTc tgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggTcag 6060  
 cgaggccaag caggccgctg tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttccttgTtc gatattgcgc cgtggccgga cagcatgca gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccc cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cagTcaaca aggacgtgaa gatcacctac accggcgTcg agctgcgggc 6300  
 cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtgcgccta caggcgacgg cgatgggctt 6480  
 cagTccgac cgcgTtgggc acctggaaTc ggtgtcgctg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgTgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgTcgccgac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgcg tcaagctgga 6720  
 aaccttccgc ctcatgtcg gatcgattc caccgcgTg aagaagtggc gcgagcaggt 6780  
 cggcgaaaggc tgcgaaagat tgcgaggcag cggcctggTg gaacacgcct gggTcaatga 6840  
 tgacctggTg cattgcaaac gctagggcct tgtggggTca gttccggctg ggggttcagc 6900  
 agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc 6960  
 tcagtatcg tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cgcccgacgt gcaggatttc 7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tggttcgggtc cgttttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgacgggt tgcgagatgc cgtggcattc 7200  
 ggcgccctaca tcgaaggcga gatcattggg ctgtcgggtc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 ggggtcggcg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaattatt 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggagg ggtagcgagg 7500  
 acggtagggc ctgtgcagcc gctgatggtc gtgttcattc ctgccgtctc gctaggtagc 7560  
 ccgatacgat tgatggcggc cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggccggggcg 7680  
 gtttccatgg cgttcgggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc 7740  
 acctttaccg cctggcaact ggcgcccgga ggactctctc tcgttccagt agcttttagt 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tgggtccggg ggatctcgcg actcgaacct 7920  
 acagttgttt ccttactggg cttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggtg agcaatggat 8040  
 aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttccctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagtcc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct ccgctgacg ccgtcccgga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggctc gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtgggt taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgccctggcc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaat atcctgtttg atgggtggtc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg 8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag cggggaacg tggcgagaaa 8940  
 ggaaggggaag aaagcgaaag gagcggggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaa 9660  
 agttcgggtg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttogcccaa tagcagccag 9900  
 tcccttcccg cttcagtac aacgtcgagc acagctgcgc aaggaacgcc cgctcgggcc 9960  
 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcacccga caggtcgggtc 10020  
 ttgacaaaaa gaaccgggcg ccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatcaa gctcccatgg gccctcgact agagtogaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcgggc 10380  
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcgatgatc ttgatocct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gttcccaac 10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 10680  
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 actggctttc tacgtgttcc gcttccttta gcagcccttg cgccctgagt gcttgcgga 10800  
 gcgtgaagct tgcattgctg caggtcgacg gcgcgcgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacy 10980  
 tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatgta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgaggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagattata atttcttcat 11520  
 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggctctctgc tttaatgaga 11580  
 tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 11640  
 aaacctgagc atgtgtagct cagatcetta ccgcggtttt cggttcattc taatgaatat 11700  
 atcaccggtt actatcgat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760  
 gtogagcaaa tttacacatt gccactaaac gtctaaaccc ttgtaatttg tttttgtttt 11820  
 actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggtact aaatttataa 11880  
 caccttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940  
 aattattttt gtcttctaaa tacatatact aatcaactgg aatgtaaat atttgcta 12000  
 atttctacta taggagaatt aaagtgtgtg aatgtgtac cacaaggttt ggagatttaa 12060  
 ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120  
 aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 12180  
 tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg 12240  
 aaagtttaaa aatatttttg aatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300

cacttggagg atgcaataat gaagaaaact acaaatttac atgcaactag ttatgcatgt 12360  
 agtctatata atgaggattt tgcaatactt tcattcatac acactcacta agttttacac 12420  
 gattataatt tcttcatagc cagcggatcc gatatcgggc ccgctagcgt taaccctgct 12480  
 ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttgtgca 12540  
 cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc gggttcattct 12600  
 aatgaatata tcaccogtta ctatcgattt tttatgaata atattctccg ttcaatttac 12660  
 tgattgtccg tcgacgaatt cgagctcggc gcgcctctag aggatcgatg aattcagatc 12720  
 ggctgagtgg ctcttcaac gttgcggttc tgtcagttcc aaacgtaaaa cggcttgtcc 12780  
 cgcgtcatcg gcgggggtca taacgtgact cccttaatto tccgctcatg atcagattgt 12840  
 cgtttccgc cttcagttta aactatcagt gtttgacagg atatattggc gggtaaacct 12900  
 aagagaaaag agcgtttatt agaataatcg gatatttaaa agggcgtgaa aaggtttatt 12960  
 cttcgtccat ttgtatgtgc atgccaacca cagggttccc ca 13002

<210> 24  
 <211> 13905  
 <212> DNA  
 <213> Unknown

<220>  
 <223> pflanzlicher Expressionsvektor mit drei  
 Promotor-Terminator-Expressionskassetten

<400> 24  
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 gcgcccagca cagggtgcga ggcaaattgc accaacgcat acagcgccag cagaatgccca 120  
 tagtgggicg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg cagcacgggc 180  
 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcaogtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420  
 gcgtagacgg tctgaogaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc 480  
 tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgtgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 ccggcacgcg accggggcga ccgcagatgg aaacggccga cgcgcagctt cgcttctct 720

gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcggca 840  
 ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc .ttcgacgaag 900  
 ccggtccgga cgcagcgctc gagcaggggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccaccgcgtc agacgcccg agcagcccg tcagggcttt ttcatgccct gccctagcgt 1140  
 ccaagcctca cggccgcgtc cggcctctct ggccggccttc tggcgctctt ccgcttcctc 1200  
 gctcactgac tcgctgcgct cggctgctcg gctgcggcga gcggtatcag ctactcaaa 1260  
 ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagagggtggc gaaaccgcac 1440  
 aggactataa agataccagg cgtttcccc tcggaagctcc ctcgctgcgt ctctgttcc 1500  
 gacctgccc cttaccgat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccatcctttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttcggtgtaga ctttcttgg tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccaccgcg gagcgggtgt tccttcttca 1740  
 ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggtaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggcgcctt gggcggcctg ctgaaactct ggctcaccga cgaccgcgc acggcgcggt 2100  
 tcggtgatgc cagatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggatcat gatggcggtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaa cagctcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgtcacc gggctggttg cctcgccgc tgggtggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgcgggc gttgtggata 2460

cctcgcggaa aacttgcccc tctactgacag atgaggggag gacgttgaca cttgagggggc 2520  
 cgactcaccc ggcgcgggct tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcctt gcgggtattga cacttgaggg gcgcgactac 2700  
 tgacagatga ggggcgggat ccttgacact tgaggggcag agtgctgaca gatgagggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggttt 2820  
 ccgcccgttt ttcggccacc gctaacctgt cttttaacct gtttttaaac caatatttat 2880  
 aaaccttggt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctogaaccc tcccggcccc ctaacgcggg cctcccatcc cccagggggc 3000  
 tgcccccctc ggccgcgaac ggccctaccc caaaaatggc agcgctggca gtcccttgcca 3060  
 ttgccgggat cggggcagta acgggatggg cgatcagccc gagcgcgacg ccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcacgcacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 gcggcctggg tggcgccctg cccttcaact cggccgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcatte ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgagtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatgtt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcagggtta tgccgctcaa ttgcgtcgt atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcacc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgctc aacaaccgtc ttccggagac tgtcatcgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc ccgcacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgcccggctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgcca taatgcgggc tgttgcccgg catccaacgc cattcatggc 4200

catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagtgtg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcgggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 tggtttcaaa atcgggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctgggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggatatata agctggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtcttttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160  
 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcaggcgga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg gggagaaca gtatgtcgag 5340  
 ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcacat agtggttttg ctctcaggcc gaggccacg gcaagtattt 5520  
 gggcaagggg tcgctgggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcac tgcgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaac aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga 5880  
 gcgcgacagc gtgcaactgg ctcccctgc cctgcccgcg ccatcggccg ccgtggagcg 5940

ttgcgctcgt ctgcaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacagggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttccttggtc gatattgcgc cgtggccgga cagcatgcga gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccc cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgctcg agctgcgggc 6300  
 cgacgatgac gaactgggtg ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tcgcgctcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgcg tcaagctgga 6720  
 aaccttcgcg ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgogaagagt tgcgaggcag cggcctgggt gaacacgcct gggccaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc 6900  
 agccagcgct ttactggcat ttcaggaaca agcgggcact gtcgacgca cttgcttcgc 6960  
 tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc 7200  
 ggcgccctaca tcgacggcga gatcattggg ctgtcggctc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 ggggtcgcgc gtatgtcgtc gcggcggttg ccggcgggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccggggcg ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggc gtgttcattc ctgccgctc gctaggtagc 7560  
 ccgatacgat tgatggcggt cctgggggct atttgcgga ctgcgggcgt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtcggcgctc cagcgggcct ggccggggcg 7680

gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccggt gcctctgctc 7740  
 acctttaccg cctggcaact ggcgcccgga ggacttctgc tcgttccagt agcttttagtg 7800  
 tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcgga cttcgggtcc ccgacctgta ccattcgggtg agcaatggat 8040  
 aggggagttg atatcgta cgttcacttc taaagaaata gcgcactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagtgc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggagcaacg ctctgtcctc gttacaatca acatgtctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccog gcagcttagt tgcggttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgcgcctta caacggctct ccgctgacg ccgtcccga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgcttgccc tgagagagtt gcagcaagcg gtccacgctg gtttgccca 8640  
 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggt tgagtgttgt tccagtttg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gtttttggg gtcgaggtgc cgtaaagcac taaatcggaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaagggag aaagcgaaag gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcggtgcg ggctctctcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaatttt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420

tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagtct 9480  
 tcagcaatat caccgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtoga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc ggccttgag cctggcgaac 9660  
 agttcggctg ggcgagccc ctgatgtct. tcgtccagat catcctgac gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgccaa tagcagccag 9900  
 tcccttcccg cttcagtgc aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctgcctcgc agttcattca gggcacccga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatacgcc aatagcctct ccaccaagc ggcggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gtcctcatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatgtgc tagctgatag tgacctagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaacgg cttgtccgc 10440  
 gtcatcggcg ggggtcataa cgtgactccc ttaattctcc gtcctatgc ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgccca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcgtt 10680  
 ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 actggctttc tacgtgttcc gcttcttcta gcagcccttg cgcctgagt gcttgccgca 10800  
 gcgtgaagct tgcatgcctg caggtcgacg gcgcgccgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg 10980  
 tttgccaaca cttagcaatt tgcaagtga ttaattgatt ctaaattatt tttgtctct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgc aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160

ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaacatgac atccacttg aggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagcattata atttcttcat 11520  
 agccagccca ccgcggtggg cggccgcctg cagtctagaa ggcctcctgc tttaatgaga 11580  
 tatgcgagac gcctatgac gcgatgatt tgctttcaat tctgttggtc acgttgtaaa 11640  
 aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 11700  
 atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtcc 11760  
 gtcgagcaaa tttacacatt gccactaac gtctaaaccc ttgtaatttg tttttgtttt 11820  
 actatgtgtg ttatgtattt gatttgcgat aaatttttat atttggtact aaatttataa 11880  
 cactttttat gctaacgttt gccaacactt agcaatttgc aagttgatta attgattcta 11940  
 aattattttt gtcttctaaa tacatatact aatcaactgg aaatgtaaat atttgctaatt 12000  
 atttctacta taggagaatt aaagtgagtg aatattgtac cacaaggttt ggagatttaa 12060  
 ttgttgcaat gctgcatgga tggcatatac accaaacatt caataattct tgaggataat 12120  
 aatggtacca cacaagattt gaggtgcatg aacgtcacgt ggacaaaagg tttagtaatt 12180  
 tttcaagaca acaatgttac cacacacaag ttttgaggtg catgcatgga tgccctgtgg 12240  
 aaagtttaaa aatatttttg aaatgatttg catggaagcc atgtgtaaaa ccatgacatc 12300  
 cacttgagg atgcaataat gaagaaaact acaaatctac atgcaactag ttatgcatgt 12360  
 agtctatata atgaggattt tgcaatactt tcattcatac aactcacta agttttacac 12420  
 gattataatt tcttcatagc cagcggatcc gatatcggtc ccgctagcgt taaccctgct 12480  
 ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt ctgttggtga 12540  
 cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgcgggttc ggttcattct 12600  
 aatgaatata tcacccgtta ctatcgtatt tttatgaata atattctccg ttcaatttac 12660  
 tgattgtccg tcgagcaaat ttacacattg ccactaaacg tctaaaccct tgtaatttgt 12720  
 ttttgtttta ctatgtgtgt tatgtatttg atttgcgata aatttttata tttggtacta 12780  
 aatttataac accttttatg ctaacgtttg ccaacttca gcaatttgca agttgattaa 12840  
 ttgattctaa attatttttg tcttctaaat acatatacta atcaactgga aatgtaaata 12900

ttgctaata tttctactat aggagaatta aagtgagtga atatggtacc acaaggtttg 12960  
 gagatttaat tgttgcaatg ctgcatggat ggcatatata ccaaaccattc aataattcctt 13020  
 gaggataata atggtaccac acaagatttg aggtgcatga acgtcacgtg gacaaaaggt 13080  
 ttagtaattt ttcaagacaa caatgttacc acacacaagt tttgaggtgc atgcatggat 13140  
 gccctgtgga aagtttaaaa atatttttga aatgatttgc atggaagcca tgtgtaaaac 13200  
 catgacatcc acttggagga tgcaataatg aagaaaacta caaatttaca tgcaactagt 13260  
 tatgcatgta gtctatataa tgaggatttt gcaatacttt cattcatata cactcactaa 13320  
 gttttacacg attataatctt ctcatagcc agcagatctg ccggcatcga tcccgggcca 13380  
 tggcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttg ctttcaattc 13440  
 tgttgtgcac gttgtaaaaa acctgagcat gtgtagctca gatccttacc gccgggttcg 13500  
 gttcattcta atgaatatat caccggttac tatcgtatctt ttatgaataa tattctccgt 13560  
 tcaatttact gattgtccgt cgacgagctc ggcgcgcctc tagaggatcg atgaattcag 13620  
 atcggctgag tggctccttc aacgttgccg ttctgtcagt tccaaacgta aaacggcttg 13680  
 tcccgctca tccggcgggg tcataacgtg actcccttaa ttctccgctc atgatcagat 13740  
 tgtcgtttcc cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa 13800  
 cctaagagaa aagagcgttt attagaataa tcggatattt aaaaggcgt gaaaagggtt 13860  
 atccttcgtc catttgtatg tgcattgcaa ccacagggtt cccca 13905

<210> 25  
 <211> 15430  
 <212> DNA  
 <213> Unknown

<220>  
 <223> pflanz. Expressionsvektor mit zwei Promotor-  
 Terminator-Expressionskassetten inseriert ist  
 Physcomitrella patens Elongase und Desaturase

<220>  
 <221> CDS  
 <222> (11543) .. (12415)

<220>  
 <221> CDS  
 <222> (13313) .. (14890)

<400> 25  
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag cagaatgcc 120

tagtgggagg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
 ataatacaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatec gtgccgccct ggacctgttg aacgaggtcg 420  
 gcgtagacgg tctgacgaca cgcaaactgg cggaaacggtt gggggttcag cagccggcgc 480  
 ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 cgggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720  
 gcgagggcggg tttttcggcc ggggacgcgc tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga 840  
 ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttgcacgaag 900  
 ccggtccgga cgcagcgctt gagcagggac tcgcggtgat tgcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caacccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080  
 ccacgcgctc agacgcccggt agcagcccgc tacgggcttt ttcatgccct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggcggccttc tggcgctctt ccgcttcctc 1200  
 gctcactgac tcgtgcgct cggctgcttc gctgcccga gcggtatcag ctactcaaa 1260  
 ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt tccataggct 1380  
 ccgccccctc gacgagcatc aaaaaaatcg acgctcaagt cagagggtggc gaaacccgac 1440  
 aggactataa agataccagg cgtttcccc tggaaagctc ctcgtagcgt ctctgttcc 1500  
 gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatctt ttccggtatat ccatcctttt 1620  
 tcgcacgata tacaggatct tgccaaaggg ttctgttaga ctttccttgg tgtatccaac 1680  
 ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgcccg cgtaacagat gagggcaagc ggatggctga tgaaaccaag ccaaccagga 1860

agggcgagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca 1980  
 aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc 2040  
 tgggcccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt 2100  
 tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat caagaagagc 2280  
 gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggctgggtg ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgccgtcga agccgtgtgc gagacaccgc ggccgccggc gttgtggata 2460  
 cctcgcgga aacttggccc tcaactgacag atgaggggcg gacgttgaca cttgaggggc 2520  
 cgactcacc ggcgcgcgct tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tcccacagat 2640  
 gatgtggaca agcctgggga taagtgcct gcggtattga cacttgagg ggcgactac 2700  
 tgacagatga ggggcgcgat ccttgacact tgaggggcag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaagggttt 2820  
 ccgcccgttt ttcgccacc gctaacctgt cttttaacct gcttttaaac caatatattat 2880  
 aaaccttgtt ttttaaccag gctgcgcct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccct tctgaaccc tcccgcccg ctaacgcggg cctcccatcc cccaggggc 3000  
 tgccccctc ggccgogaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccgggat cggggcagta acgggatgg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcacgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 gcggcctggg tggcggcctg cccttcactt cggcgtcgg ggcattcacg gacttcatgg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgaggtg ataggtaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cctttacaga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgaaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600

tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccatgactt 3720  
 tgcattgcag ctccaccgat ttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcaggta tgccgctcaa ttccgtcgt atatcgctt ctgattacgt 3840  
 gcagctttcc ctccaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt ccccatagt cgttcaccga 3960  
 atacgtgogc aacaaccgtc ttccggagac tgcataacg gtaaacagc cagcgctggc 4020  
 gcgatttagc ccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgcccggtg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgcca taatcgggc tgttgcccg catccaacg cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgtgatgt ccggcggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca cccataattg 4440  
 tggtttcaaa atcggtccg tcgatactat gttatacgcc aactttgaaa acaactttga 4500  
 aaaagctgtt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagttcgt 4560  
 cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aaggatatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt ccttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtcttttc actccatcga catatcggt tgcctcata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg attttttaaa gacggaaaag 5160  
 cccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcaggcgga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg ggaagaaca gtatgtcgag 5340

ctatTTTTTg acttactggg gatcaagcct gattgggaga aaataaaata ttatatttta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcatca agtgTTTTgg ctctcaggcc gaggcccacg gcaagtattt 5520  
 gggcaagggg tcgctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcac tgccgataag gtggattatc tggacaccaa 5640  
 ggcaccaggc ggggtcaaatac aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 cccgcaagga ggggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttccgccc aggatgccga aaccatcgca agccgcaccg tcatgctgac 5820  
 gccccgcgaa accttcacgt ccgtcggctc gatgggtccag caagctacgg ccaagatcga 5880  
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccacggcccg ccgtggagcg 5940  
 ttcgctcgt ctcgaaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgctg tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttcttgttc gatattgcgc cgtggccgga cagcatcgca gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa 6240  
 ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc 6300  
 cgacgatgac gaactgggtg ggcagcaggc gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtgcgccta caggcgacgg cgatgggctt 6480  
 cacgtccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct 6600  
 gtttctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aacctccgc ctcatgtgcg gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcaagagt tgcgaggcag cggcctggtg gaacacgcct ggggtcaatga 6840  
 tgacctggtg cattgcaaac gctaggcct tgtgggtca gttccggctg ggggttcagc 6900  
 agccagcgt ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc 6960  
 tcagtatcgc tggggacgca cggcgcgtc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080

cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tggtcgggtc cgtttacgag 7140  
 cacgaggaga aaaagcccat ggaggcgttc gctgaacggg tgcgagatgc cgtggcattc 7200  
 ggcgcctaca tcgacggcga gatcattggg ctgtcgggtc tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaaca gcgaggccga 7320  
 ggggtcgcgc gtatgtcgtc gcgggcgttg ccggcgggtt tattgtcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcattctca tcctcggcgc acttaatat 7440  
 tcgctattct ggagcttggt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggc gtgttcattc ctgccgtct gctaggtagc 7560  
 ccgatacgat tgatggcggc cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg 7620  
 gtgttgacac caaacgcagc gctagatcct gtccgcgtcg cagcgggcct ggccggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctccgt gcctctgctc 7740  
 acctttaccg cctggcaact ggccggcgga ggacttctgc tcgttcagc agctttagt 7800  
 tttgatccgc caatccgat gcctacagga accaatgttc tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttctt tggttcggg ggatctcgc actcgaacct 7920  
 acagttgttt ccttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcggaa cttcgggtcc ccgacctgta ccattcggg agcaatggat 8040  
 aggggagttg atatcgtaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagtcc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggagcaacg ctctgtcacc gttacaatca acatgctacc ctccgcgaga 8280  
 tcattccgtg ttcaaaccgc gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccgga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtgtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtgggttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggt tgagtgtgt tccagtttg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat caggcgatg gccactacg 8820

tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa 8940  
 ggaaggggaag aaagcgaaag gagcggggcg cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcgggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagt 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gcgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaaagctct 9480  
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc ggccttgag cctggcgaa 9660  
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc gaatgggcag 9780  
 gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 tcccttcccg cttcagtga acgctcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctgcctctgc agttcattca gggcacccga caggtcggct 10020  
 ttgacaaaaa gaaccggggc cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc ggccggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacgatc ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560

cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcccc	10620
gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt	10680
ttcccttgtc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg	10740
actggctttc tacgtgttcc gtttcttta gcagcccttg cgccctgagt gcttgcgcca	10800
gcgtgaagct tgcatgcctg caggctgacg gcgcgcggag ctccctcgagc aaatttacac	10860
attgccacta aacgtctaaa cccttgtaat ttgtttttgt tttactatgt gtgttatgta	10920
tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacg	10980
tttgccaaca cttagcaatt tgcaagttga ttaattgatt ctaaattatt tttgtcttct	11040
aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga	11100
attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat	11160
ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga	11220
tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt	11280
taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt	11340
tggaaatgat ttgcatggaa gccatgtgta aaaccatgac atccacttgg aggatgcaat	11400
aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataattgagga	11460
ttttgcaata ctttcattca tacacactca ctaagtttta cagcattata atttcttcac	11520
agccagcccc ccgcggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag	11572
Met Glu Val Val Glu Arg Phe Tyr Gly Glu	
1 5 10	
ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt	11620
Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe	
15 20 25	
ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctg gtt	11668
Gly Val Glu Leu Thr Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val	
30 35 40	
gac agt ccc aca ccc atc gtc ctg ggt gtt tct gta tac ttg act att	11716
Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile	
45 50 55	
gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc	11764
Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg	
60 65 70	
gcc tcg gag cca ttt ttg ctg caa gct ttg gtg ctt gtg cac aac ctg	11812
Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu	
75 80 85 90	
ttc tgt ttt gcg ctg agt ctg tat atg tgc gtg ggc atc gct tat cag	11860

Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln	
95 100 105	
gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa	11908
Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys	
110 115 120	
cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac	11956
His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr	
125 130 135	
gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg	12004
Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg	
140 145 150	
caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att	12052
Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Ser Leu Ile	
155 160 165 170	
tgg tgg gct att gct cat cac gct cct ggc ggt gaa gca tat tgg tct	12100
Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser	
175 180 185	
gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc	12148
Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe	
190 195 200	
ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt	12196
Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu	
205 210 215	
ttt tgg ggc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg	12244
Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu	
220 225 230	
aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca	12292
Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro	
235 240 245 250	
caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt	12340
Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe	
255 260 265	
ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga	12388
Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly	
270 275 280	
aag caa aag gga gct aaa act gag tga tctagaaggc ctctgtctt	12435
Lys Gln Lys Gly Ala Lys Thr Glu	
285 290	
aatgagatat gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg	12495
ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa	12555
tgaatatatc acccgttact atcgtatttt tatgaataat attctccggt caatttactg	12615
attgtccgtc gagcaaatat acacattgcc actaaacgtc taaacccttg taatttgttt	12675

ttgttttact atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tggtaactaaa	12735
tttataacac cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt	12795
gattctaaat tatttttgtc ttctaaatac atatactaata caactggaaa tgtaaatatt	12855
tgctaataatt tctactatag gagaattaaa gtgagtgaat atggtagcac aagggttgga	12915
gatttaattg ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga	12975
ggataataat ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaagggtt	13035
agtaattttt caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc	13095
cctgtggaaa gtttaaaaat attttggaaa tgatttgcata ggaagccatg tgtaaaacca	13155
tgacatccac ttggaggatg caataatgaa gaaaactaca aatttcatg caactagtta	13215
tgcatgtagt ctatataatg aggattttgc aatactttca ttcatacaca ctactaagt	13275
tttacacgat tataattttc tcatagccag cggatcc atg gta ttc gcg ggc ggt	13330
Met Val Phe Ala Gly Gly	295
gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att	13378
Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile	
300 305 310	
gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act	13426
Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr	
315 320 325	
gtt ggt tcg tgg agc gta cac agt ata caa cct ttg aag cgc ctg acg	13474
Val Gly Ser Trp Ser Val His Ser Ile Gln Pro Leu Lys Arg Leu Thr	
330 335 340 345	
agt aag aag cgt gtt tcg gaa agc gct gcc gtg caa tgt ata tca gct	13522
Ser Lys Lys Arg Val Ser Glu Ser Ala Ala Val Gln Cys Ile Ser Ala	
350 355 360	
gaa gtt cag aga aat tcg agt acc cag gga act gcg gag gca ctc gca	13570
Glu Val Gln Arg Asn Ser Ser Thr Gln Gly Thr Ala Glu Ala Leu Ala	
365 370 375	
gaa tca gtc gtg aag ccc acg aga cga agg tca tct cag tgg aag aag	13618
Glu Ser Val Val Lys Pro Thr Arg Arg Ser Ser Gln Trp Lys Lys	
380 385 390	
tcg aca cac ccc cta tca gaa gta gca gta cac aac aag cca agc gat	13666
Ser Thr His Pro Leu Ser Glu Val Ala Val His Asn Lys Pro Ser Asp	
395 400 405	
tgc tgg att gtt gta aaa aac aag gtg tat gat gtt tcc aat ttt gcg	13714
Cys Trp Ile Val Val Lys Asn Lys Val Tyr Asp Val Ser Asn Phe Ala	
410 415 420 425	
gac gag cat ccc gga gga tca gtt att agt act tat ttt gga cga gac	13762

Asp	Glu	His	Pro	Gly	Gly	Ser	Val	Ile	Ser	Thr	Tyr	Phe	Gly	Arg	Asp		
				430					435					440			
ggc	aca	gat	gtt	ttc	tct	agt	ttt	cat	gca	gct	tct	aca	tgg	aaa	att	13810	
Gly	Thr	Asp	Val	Phe	Ser	Ser	Phe	His	Ala	Ala	Ser	Thr	Trp	Lys	Ile		
			445					450					455				
ctt	caa	gac	ttt	tac	att	ggt	gac	gtg	gag	agg	gtg	gag	ccg	act	cca	13858	
Leu	Gln	Asp	Phe	Tyr	Ile	Gly	Asp	Val	Glu	Arg	Val	Glu	Pro	Thr	Pro		
		460					465					470					
gag	ctg	ctg	aaa	gat	ttc	cga	gaa	atg	aga	gct	ctt	ttc	ctg	agg	gag	13906	
Glu	Leu	Leu	Lys	Asp	Phe	Arg	Glu	Met	Arg	Ala	Leu	Phe	Leu	Arg	Glu		
	475					480					485						
caa	ctt	ttc	aaa	agt	tcg	aaa	ttg	tac	tat	gtt	atg	aag	ctg	ctc	acg	13954	
Gln	Leu	Phe	Lys	Ser	Ser	Lys	Leu	Tyr	Tyr	Val	Met	Lys	Leu	Leu	Thr		
490					495					500					505		
aat	gtt	gct	att	ttt	gct	gcg	agc	att	gca	ata	ata	tgt	tgg	agc	aag	14002	
Asn	Val	Ala	Ile	Phe	Ala	Ala	Ser	Ile	Ala	Ile	Ile	Cys	Trp	Ser	Lys		
				510					515					520			
act	att	tca	gcg	gtt	ttg	gct	tca	gct	tgt	atg	atg	gct	ctg	tgt	ttc	14050	
Thr	Ile	Ser	Ala	Val	Leu	Ala	Ser	Ala	Cys	Met	Met	Ala	Leu	Cys	Phe		
			525					530					535				
caa	cag	tgc	gga	tgg	cta	tcc	cat	gat	ttt	ctc	cac	aat	cag	gtg	ttt	14098	
Gln	Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe	Leu	His	Asn	Gln	Val	Phe		
		540					545					550					
gag	aca	cgc	tgg	ctt	aat	gaa	gtt	gtc	ggg	tat	gtg	atc	ggc	aac	gcc	14146	
Glu	Thr	Arg	Trp	Leu	Asn	Glu	Val	Val	Gly	Tyr	Val	Ile	Gly	Asn	Ala		
	555					560					565						
gtt	ctg	ggg	ttt	agt	aca	ggg	tgg	tgg	aag	gag	aag	cat	aac	ctt	cat	14194	
Val	Leu	Gly	Phe	Ser	Thr	Gly	Trp	Trp	Lys	Glu	Lys	His	Asn	Leu	His		
570					575					580					585		
cat	gct	gct	cca	aat	gaa	tgc	gat	cag	act	tac	caa	cca	att	gat	gaa	14242	
His	Ala	Ala	Pro	Asn	Glu	Cys	Asp	Gln	Thr	Tyr	Gln	Pro	Ile	Asp	Glu		
				590				595						600			
gat	att	gat	act	ctc	ccc	ctc	att	gcc	tgg	agc	aag	gac	ata	ctg	gcc	14290	
Asp	Ile	Asp	Thr	Leu	Pro	Leu	Ile	Ala	Trp	Ser	Lys	Asp	Ile	Leu	Ala		
			605					610						615			
aca	gtt	gag	aat	aag	aca	ttc	ttg	cga	atc	ctc	caa	tac	cag	cat	ctg	14338	
Thr	Val	Glu	Asn	Lys	Thr	Phe	Leu	Arg	Ile	Leu	Gln	Tyr	Gln	His	Leu		
		620					625					630					
ttc	ttc	atg	ggt	ctg	tta	ttt	ttc	gcc	cgt	ggt	agt	tgg	ctc	ttt	tgg	14386	
Phe	Phe	Met	Gly	Leu	Leu	Phe	Phe	Ala	Arg	Gly	Ser	Trp	Leu	Phe	Trp		
	635					640					645						
agc	tgg	aga	tat	acc	tct	aca	gca	gtg	ctc	tca	cct	gtc	gac	agg	ttg	14434	
Ser	Trp	Arg	Tyr	Thr	Ser	Thr	Ala	Val	Leu	Ser	Pro	Val	Asp	Arg	Leu		
650					655					660					665		

ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca	14482
Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr	
670 675 680	
gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg	14530
Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val	
685 690 695	
act gag ctc atg tcc ggc atg ctg ctg ggc ttt gta ttt gta ctt agc	14578
Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser	
700 705 710	
cac aat ggg atg gag gtt tat aat tgc tct aaa gaa ttc gtg agt gca	14626
His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala	
715 720 725	
cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg	14674
Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp	
730 735 740 745	
ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca	14722
Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr	
750 755 760	
atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc	14770
Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe	
765 770 775	
tgt aag aaa cac ggt ctg gtg tac gaa gac gta tct att gct acc ggc	14818
Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly	
780 785 790	
act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca	14866
Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala	
795 800 805	
gag cag cat gct acc acc agt taa gctagcgtta accctgcttt aatgagatat	14920
Glu Gln His Ala Thr Thr Ser	
810 815	
gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa	14980
cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc	15040
accggttact atcgtatttt tatgaataat attctccgtt caatttactg attgtccgtc	15100
gaogaattcg agctcggcgc gcctctagag gatcgatgaa ttcagatcgg ctgagtggct	15160
ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg gcttgtcccg cgtcatcggc	15220
gggggtcata acgtgactcc cttaattctc cgctcatgat cagattgtcg tttcccgctt	15280
tcagtttaaa ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag	15340
cgtttattag aataatcgga tattttaaag ggcgtgaaaa ggtttatcct tcgtccattt	15400
gtatgtgcat gcccaaccaca gggttcccca	15430

&lt;210&gt; 26

&lt;211&gt; 290

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;400&gt; 26

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser  
 1 5 10 15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp  
 20 25 30

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
 35 40 45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu  
 50 55 60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu  
 65 70 75 80

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
 85 90 95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr  
 100 105 110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile  
 115 120 125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr  
 130 135 140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His  
 145 150 155 160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His  
 165 170 175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly  
 180 185 190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg  
 195 200 205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
 210 215 220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr  
 225 230 235 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
 245 250 255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr  
 260 265 270

Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
 275 280 285

Thr Glu  
 290

<210> 27  
 <211> 525  
 <212> PRT  
 <213> Unknown

<400> 27  
 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
 1 5 10 15

Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
 20 25 30

Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
 35 40 45

Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
 50 55 60

Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
 65 70 75 80

Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
 85 90 95

Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
 100 105 110

His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
 115 120 125

Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
 130 135 140

Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
 145 150 155 160

Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
 165 170 175

Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
 180 185 190

Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
 195 200 205

Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
 210 215 220

Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys  
 225 230 235 240

Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe  
 245 250 255  
 Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly  
 260 265 270  
 Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys  
 275 280 285  
 Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr  
 290 295 300  
 Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp  
 305 310 315 320  
 Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile  
 325 330 335  
 Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg  
 340 345 350  
 Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu  
 355 360 365  
 Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr  
 370 375 380  
 Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro  
 385 390 395 400  
 Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly  
 405 410 415  
 Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser  
 420 425 430  
 Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly  
 435 440 445  
 Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu  
 450 455 460  
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala  
 465 470 475 480  
 Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp  
 485 490 495  
 Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu  
 500 505 510  
 Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser  
 515 520 525

&lt;210&gt; 28

&lt;211&gt; 17752

<212> DNA  
 <213> Unknown

<220>  
 <223> pflanz. Expressionsvektor mit 3  
 Promotor-Terminator- Expressionskassetten  
 inseriert mit Physcomitrella Elongase + Desaturase  
 + Phaeodactylum Desaturase

<220>  
 <221> CDS  
 <222> (11543)..(12415)

<220>  
 <221> CDS  
 <222> (13313)..(14890)

<220>  
 <221> CDS  
 <222> (15791)..(17200)

<400> 28  
 gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg atccgacagc 60  
 ggcgccagca caggtgcgca ggcaaattgc accaacgcac acagcgccag cagaatgccca 120  
 tagtggggcg tgacgtcggt cgagtgaacc agatcgcgca ggaggcccg cagcaccggc 180  
 ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac gatcaggggt 240  
 atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg aacgcgcgga 300  
 ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa gtgtcaagca 360  
 tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg aacgaggtcg 420  
 gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt ggggggttcag cagccggcgc 480  
 ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa gccatgctgg 540  
 cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca tttctgatcg 600  
 ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg cgcattccatg 660  
 ccggcacgcg accggcgca ccgcagatgg aaacggccga cgcgcagctt cgcttcctct 720  
 gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc agctacttca 780  
 ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag cgcggcgga 840  
 ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc ttcgacgaag 900  
 ccggtccgga cgcagcgctc gagcagggac tcgcggtgat tgtcgatgga ttggcgaaaa 960  
 ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga tcaggaccgc 1020  
 tgccggagcg caaccactc actacagcag agccatgtag acaacatccc ctcccccttt 1080

ccaccgcgtc agacgcccggt agcagcccgc tacgggcttt ttcattgccct gccctagcgt 1140  
 ccaagcctca cggccgcgct cggcctctct ggccgcttc tggcgctctt ccgcttcctc 1200  
 gctcactgac tcgctgcgct cggtcgttcg gctcggcgga gcggtatcag ctcactcaaa 1260  
 ggccgtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa 1320  
 aggccagcaa aaggccagga accgtaaaaa ggccgcttg ctggcgcttt tccataggct 1380  
 ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac 1440  
 aggactataa agataccagg cgtttcccc tggaaagctc ctcgtgcgct ctcctgttcc 1500  
 gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttt 1560  
 ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat ccctcctttt 1620  
 tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg tgtatccaac 1680  
 ggcgtagcc ggccaggata ggtgaagtag gccacccgc gagcgggtgt tccttcttca 1740  
 ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga ggctggccgg 1800  
 ctaccgccgg cgtaacagat gagggaagc ggatggctga tgaaaccaag ccaaccagga 1860  
 agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa 1920  
 agggcgccgc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggttaca 1980  
 aatcacggg cgtcgtggac tatgagcag tccgcgagct ggccgcac caatggcgacc 2040  
 tgggcccctt ggccggcctg ctgaaactct ggctcaccga cgaccgcgc accggcgcggt 2100  
 tcggtgatgc caccatctc gccctgctgg cgaagatcga agagaagcag gacgagcttg 2160  
 gcaaggatcat gatggcgctg gtccgcccga gggcagagcc atgacttttt tagccgctaa 2220  
 aacggccggg gggtgcgctg gattgccaag cagctcccca tgcgctccat caagaagagc 2280  
 gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga gcgcctttgc 2340  
 gacgctcacc gggttggttg ccctgcgcgc tgggctggcg gccgtctatg gccctgcaaa 2400  
 cgcgccagaa acgcgctcga agccgtgtgc gagacaccgc ggccgcccgc gttgtggata 2460  
 cctcgcggaa aacttgcccc tcaactgacag atgaggggag gacgttgaca cttgaggggc 2520  
 cgactcacc ggccggcggt tgacagatga ggggcaggct cgatttcggc cggcgacgtg 2580  
 gagctggcca gcctcgcaaa tcggcgaaaa cgctgattt tacgcgagtt tccacagat 2640  
 gatgtggaca agcctgggga taagtgcct gcggtattga cacttgagg gcgcgactac 2700  
 tgacagatga gggcgcgat ccttgacact tgaggggag agtgctgaca gatgaggggc 2760  
 gcacctattg acatttgagg ggctgtccac aggcagaaaa tccagcattt gcaaggggtt 2820

ccgccccgttt ttccggccacc gctaacctgt cttttaacct gctttttaaac caatatttat 2880  
 aaaccttggt ttaaccagg gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg 2940  
 tgccccccct tctgaaccc tcccgccccg ctaacgcggg cctcccatcc cccagggggc 3000  
 tgcgccccctc ggccgcgaac ggcctcacc caaaaatggc agcgctggca gtccttgcca 3060  
 ttgccccgat cggggcagta acgggatggg cgatcagccc gagcgcgacg cccggaagca 3120  
 ttgacgtgcc gcaggtgctg gcatcgacat tcagcgacca ggtgccgggc agtgagggcg 3180  
 ggggcctggg tgggggcctg cccttcactt cggccgtcgg ggcattcacg gacttcattg 3240  
 cggggccggc aatttttacc ttgggcattc ttggcatagt ggtcgcgggt gccgtgctcg 3300  
 tgttcggggg tgcgataaac ccagcgaacc atttgagggt ataggttaaga ttataccgag 3360  
 gtatgaaaac gagaattgga cetttaacga attactctat gaagcgccat atttaaaaag 3420  
 ctaccaagac gaagaggatg aagaggatga ggaggcagat tgccttgaat atattgacaa 3480  
 tactgataag ataatatatc ttttatatag aagatatcgc cgtatgtaag gatttcaggg 3540  
 ggcaaggcat aggcagcgcg cttatcaata tatctataga atgggcaaag cataaaaact 3600  
 tgcattggact aatgcttgaa acccaggaca ataaccttat agcttgtaaa ttctatcata 3660  
 attgggtaat gactccaact tattgatagt gttttatggt cagataatgc ccgatgactt 3720  
 tgtcatgcag ctccaccgat tttgagaacg acagcgactt ccgtcccagc cgtgccaggt 3780  
 gctgcctcag attcaggtta tgccgctcaa ttcgctgctg atatcgcttg ctgattacgt 3840  
 gcagctttcc cttcaggcgg gattcataca gcggccagcc atccgtcatc catatcacca 3900  
 cgtcaaaggg tgacagcagg ctcataagac gcccagcgt cgccatagtg cgttcaccga 3960  
 atacgtgcgc aacaaccgtc ttccggagac tgcatacgc gtaaaacagc cagcgctggc 4020  
 gcgatttagc cccgacatag cccactgtt cgtccatttc cgcgcagacg atgacgtcac 4080  
 tgccccgctg tatgcgcgag gttaccgact gcggcctgag ttttttaagt gacgtaaaat 4140  
 cgtgttgagg ccaacgccc taatgcgggc tgttgccgg catccaacgc cattcatggc 4200  
 catatcaatg attttctggt gcgtaccggg ttgagaagcg gtgtaagtga actgcagttg 4260  
 ccatgtttta cggcagtgag agcagagata gcgctgatgt ccggcggtgc ttttgccgtt 4320  
 acgcaccacc ccgtcagtag ctgaacagga gggacagctg atagacacag aagccactgg 4380  
 agcacctcaa aaacaccatc atacactaaa tcagtaagtt ggcagcatca ccataattg 4440  
 tggtttcaaa atcggctccg tcgatactat gttatagcc aactttgaaa acaactttga 4500  
 aaaagctggt ttctggtatt taaggtttta gaatgcaagg aacagtgaat tggagtctgt 4560

cttgttataa ttagcttctt ggggtatctt taaatactgt agaaaagagg aaggaaataa 4620  
 taaatggcta aaatgagaat atcaccggaa ttgaaaaaac tgatcgaaaa ataccgctgc 4680  
 gtaaaagata cggaaggaat gtctcctgct aagggtatata agctgggtggg agaaaatgaa 4740  
 aacctatatt taaaatgac ggacagccgg tataaaggga ccacctatga tgtggaacgg 4800  
 gaaaaggaca tgatgctatg gctggaagga aagctgcctg ttccaaaggt cctgcacttt 4860  
 gaacggcatg atggctggag caatctgctc atgagtgagg ccgatggcgt cctttgctcg 4920  
 gaagagtatg aagatgaaca aagccctgaa aagattatcg agctgtatgc ggagtgcac 4980  
 aggtctcttc actccatcga catatcggat tgtccctata cgaatagctt agacagccgc 5040  
 ttagccgaat tggattactt actgaataac gatctggccg atgtggattg cgaaaactgg 5100  
 gaagaagaca ctccatttaa agatccgcgc gagctgtatg atttttttaa gacggaaaag 5160  
 ccgaagagg aacttgtctt ttcccacggc gacctgggag acagcaacat ctttgtgaaa 5220  
 gatggcaaag taagtggctt tattgatctt gggagaagcg gcaggggcga caagtggat 5280  
 gacattgcct tctgcgtccg gtcgatcagg gaggatatcg ggaagaaca gtatgtcgag 5340  
 ctattttttg acttactggg gatcaagcct gattgggaga aaataaaata ttatatatta 5400  
 ctggatgaat tgttttagta cctagatgtg gcgcaacgat gccggcgaca agcaggagcg 5460  
 caccgacttc ttccgcatca agtggttttg ctctcaggcc gagggccacg gcaagtattt 5520  
 gggcaagggg tcgctgggat tcgtgcagg caagattcgg aataccaagt acgagaagga 5580  
 cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa 5640  
 ggcaaccaggc gggtaaatac aggaataagg gcacattgcc ccggcgtgag tcggggcaat 5700  
 ccgcaagga gggatgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat 5760  
 cgacgcgggg ttttcgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc 5820  
 gccccgcgaa accttcagat ccgtcggctc gatgggtccag caagctacgg ccaagatcga 5880  
 gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccacggccg ccgtggagcg 5940  
 ttgcgctcgt ctgaaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg 6000  
 aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag 6060  
 cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct 6120  
 ttccttgctc gatattgcgc cgtggccgga cagatgcga gcgatgcaa acgacacggc 6180  
 ccgctctgcc ctgttcacca cgcgcaaca gaaaatcccg cgcgaggcgc tgcaaaaca 6240  
 ggtcattttc cagtcacaag aggcgtgaa gatcacctac accggcgtcg agctgcgggc 6300

cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg 6360  
 cgagccgatac accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg 6420  
 ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt 6480  
 cactccgac cgcgttgggc acctggaatc ggtgtcgtg ctgcaccgct tccgcgtcct 6540  
 ggaccgtggc aagaaaacgt cccgttgcca ggtcctgac gacgaggaaa tcgtcgtgct 6600  
 gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac 6660  
 ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtaccgc tcaagctgga 6720  
 aaccttcgc ctcatgtgc gatcggattc caccgcgtg aagaagtggc gcgagcaggt 6780  
 cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggatcaatga 6840  
 tgacctggtg cattgcaaac gctagggcct tgtgggtca gttccggctg ggggttcagc 6900  
 agccagcgt tttactggcat ttcaggaaca agcgggact gtcgacgca cttgcttcgc 6960  
 tcagtatgc tcgggacgca cggcgcgtc tacgaactgc cgataaacag aggattaaaa 7020  
 ttgacaattg tgattaaggc tcagattcga cggcttgag cggccgacgt gcaggatttc 7080  
 cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag 7140  
 cactaggaga aaaagcccat ggaggcgtt gctgaacggt tgcgagatgc cgtggcattc 7200  
 ggcgcctaca tcgacggcga gatcattggg ctgtcgtct tcaaacagga ggacggcccc 7260  
 aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga 7320  
 ggggtcgcg gtatgtcgt gcgggcgtt cggcggtt tattgctcgt gatgatcgtc 7380  
 cgacagattc caacgggaat ctggtggatg cgcatttca tcctcggcg acttaatt 7440  
 tcgtattct ggagcttgtt gtttatttcg gtctaccgc tgcggggcg ggtcgcggcg 7500  
 acggtaggcg ctgtgcagcc gctgatggc gtgttcatt ctgccgtct gctaggtagc 7560  
 ccgatacgat tgatggcgg cctgggggct atttgcgaa ctgcgggcgt ggcgctgtt 7620  
 gtgttgacac caaacgcagc gctagatcct gtcggcgtc cagcgggcct ggccggggcg 7680  
 gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc 7740  
 acctttaccg cctggcaact ggcgccgga ggacttctgc tcgttcagc agctttagt 7800  
 tttgatccgc caatccgat gcctacagga accaatgtt tcggcctggc gtggctcggc 7860  
 ctgatcggag cgggtttaac ctacttctt tggttccggg ggatctcgc actcgaacct 7920  
 acagttgtt cttactggg ctttctcagc ccagatctg gggtcgatca gccggggatg 7980  
 catcaggccg acagtcgga cttcgggtcc ccgacctga ccattcgtg agcaatggat 8040

aggggagttg atatcgtcaa cgttcacttc taaagaaata gcgccactca gcttcctcag 8100  
 cggctttatc cagcgatttc ctattatgtc ggcatagttc tcaagatcga cagcctgtca 8160  
 cggttaagcg agaaatgaat aagaaggctg ataattcgga tctctgcgag ggagatgata 8220  
 tttgatcaca ggcagcaacg ctctgtcatc gttacaatca acatgctacc ctccgcgaga 8280  
 tcatccgtgt ttcaaaccgg gcagcttagt tgccgttctt ccgaatagca tcggtaacat 8340  
 gagcaaagtc tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct 8400  
 gcctgtatcg agtggtgatt ttgtgccgag ctgccggtcg gggagctgtt ggctggctgg 8460  
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 8520  
 gacgttttta atgtactggg gtggtttttc ttttcaccag tgagacgggc aacagctgat 8580  
 tgcccttcac cgctggccc tgagagagtt gcagcaagcg gtccacgctg gtttgcccca 8640  
 gcaggcgaaa atcctgtttg atggtggttc cgaaatcggc aaaatccctt ataaatcaaa 8700  
 agaatagccc gagatagggg tgagtgttgt tccagtttgg aacaagagtc cactattaaa 8760  
 gaacgtggac tccaacgtca aaggcgaaa aaccgtctat cagggcgatg gccactacg 8820  
 tgaaccatca cccaaatcaa gttttttggg gtcgaggtgc cgtaaagcac taaatcgaa 8880  
 ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaaag tggcgagaaa 8940  
 ggaaggggaag aaagcgaaa gagcgggcgc cattcaggct gcgcaactgt tgggaagggc 9000  
 gatcggtgcg ggctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc 9060  
 gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 9120  
 aattaattcc catcttgaaa gaaatatagt ttaaataattt attgataaaa taacaagtca 9180  
 ggtattatag tccaagcaaa aacataaatt tattgatgca agtttaaatt cagaaatatt 9240  
 tcaataactg attatatcag ctggtacatt gccgtagatg aaagactgag tgcgatatta 9300  
 tgtgtaatac ataaattgat gatatagcta gcttagctca tcgggggatc cgtcgaagct 9360  
 agcttgggtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 9420  
 tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc gccaaactct 9480  
 tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 9540  
 ccacagtcca tgaatccaga aaagcggcca tttccacca tgatattcgg caagcaggca 9600  
 tcgccatggg tcacgacgag atcctcgccg tcgggcatgc gcgccttgag cctggcgaac 9660  
 agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgacg gacaagaccg 9720  
 gcttccatcc gagtacgtgc tcgtcgatg cgatgtttcg cttggtggtc gaatgggcag 9780

gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 9840  
 gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 9900  
 tcccttccccg cttcagtgc acgctcgagc acagctgcgc aaggaacgcc cgtcgtggcc 9960  
 agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga caggtcggtc 10020  
 ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 10080  
 ccgattgtct gttgtgcca gtcatacgcg aatagcctct ccaccaagc ggcggagaa 10140  
 cctgcgtgca atccatcttg ttcaatccaa gctcccatgg gccctcgact agagtcgaga 10200  
 tctggattga gagtgaatat gagactctaa ttggataccg aggggaattt atggaacgtc 10260  
 agtggagcat ttttgacaag aaatatttgc tagctgatag tgaccttagg cgacttttga 10320  
 acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga aaccgcggc 10380  
 tgagtggctc cttcaacgtt gcggttctgt cagttccaaa cgtaaacgg cttgtccgc 10440  
 gtcacggcg ggggtcataa cgtgactccc ttaattctcc gtcacgatc ttgatccct 10500  
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 10560  
 cttaccagag ggcgccccag ctggcaattc cggttcgctt gctgtccata aaaccgcca 10620  
 gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcgct 10680  
 ttcccttctc cagatagccc agtagctgac attcatccgg ggtcagcacc gtttctgcgg 10740  
 actggcttcc tacgtgttcc gcttcttta gcagccctg cgccctgagt gcttggcgca 10800  
 gcgtgaagct tgcatgcctg caggctcgac gcgcgcgag ctctcgagc aaatttacac 10860  
 attgccacta aacgtctaaa cccttgtaat ttgttttgt tttactatgt gtgttatgta 10920  
 tttgatttgc gataaatttt tatatttggg actaaattta taacaccttt tatgctaacy 10980  
 tttgccaaca cttagcaatt tgcaagtga ttaattgatt ctaaattatt tttgtcttct 11040  
 aaatacatat actaatcaac tggaaatgta aatatttgct aatatttcta ctataggaga 11100  
 attaaagtga gtgaatatgg taccacaagg tttggagatt taattgttgc aatgctgcat 11160  
 ggatggcata tacaccaaac attcaataat tcttgaggat aataatggta ccacacaaga 11220  
 tttgaggtgc atgaacgtca cgtggacaaa aggttttagta atttttcaag acaacaatgt 11280  
 taccacacac aagttttgag gtgcatgcat ggatgccctg tggaaagttt aaaaatattt 11340  
 tggaaatgat ttgcatggaa gccatgtgta aaacatgac atccacttg aggatgcaat 11400  
 aatgaagaaa actacaaatt tacatgcaac tagttatgca tgtagtctat ataatgagga 11460  
 ttttgcaata ctttcattca tacacactca ctaagtttta cagattata atttcttcat 11520

agccagccca ccgcggtgga aa atg gag gtc gtg gag aga ttc tac ggt gag 11572  
Met Glu Val Val Glu Arg Phe Tyr Gly Glu  
1 5 10

ttg gat ggg aag gtc tcg cag ggc gtg aat gca ttg ctg ggt agt ttt 11620  
Leu Asp Gly Lys Val Ser Gln Gly Val Asn Ala Leu Leu Gly Ser Phe  
15 20 25

ggg gtg gag ttg acg gat acg ccc act acc aaa ggc ttg ccc ctc gtt 11668  
Gly Val Glu Leu Thr Asp Thr Pro Thr Thr Lys Gly Leu Pro Leu Val  
30 35 40

gac agt ccc aca ccc atc gtc ctc ggt gtt tct gta tac ttg act att 11716  
Asp Ser Pro Thr Pro Ile Val Leu Gly Val Ser Val Tyr Leu Thr Ile  
45 50 55

gtc att gga ggg ctt ttg tgg ata aag gcc agg gat ctg aaa ccg cgc 11764  
Val Ile Gly Gly Leu Leu Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg  
60 65 70

gcc tcg gag cca ttt ttg ctc caa gct ttg gtg ctt gtg cac aac ctg 11812  
Ala Ser Glu Pro Phe Leu Leu Gln Ala Leu Val Leu Val His Asn Leu  
75 80 85 90

ttc tgt ttt gcg ctc agt ctg tat atg tgc gtg ggc atc gct tat cag 11860  
Phe Cys Phe Ala Leu Ser Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln  
95 100 105

gct att acc tgg cgg tac tct ctc tgg ggc aat gca tac aat cct aaa 11908  
Ala Ile Thr Trp Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys  
110 115 120

cat aaa gag atg gcg att ctg gta tac ttg ttc tac atg tct aag tac 11956  
His Lys Glu Met Ala Ile Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr  
125 130 135

gtg gaa ttc atg gat acc gtt atc atg ata ctg aag cgc agc acc agg 12004  
Val Glu Phe Met Asp Thr Val Ile Met Ile Leu Lys Arg Ser Thr Arg  
140 145 150

caa ata agc ttc ctc cac gtt tat cat cat tct tca att tcc ctc att 12052  
Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Ser Leu Ile  
155 160 165 170

tgg tgg gct att gct cat cac gct cct gcc ggt gaa gca tat tgg tct 12100  
Trp Trp Ala Ile Ala His His Ala Pro Gly Gly Glu Ala Tyr Trp Ser  
175 180 185

gcg gct ctg aac tca gga gtg cat gtt ctc atg tat gcg tat tac ttc 12148  
Ala Ala Leu Asn Ser Gly Val His Val Leu Met Tyr Ala Tyr Tyr Phe  
190 195 200

ttg gct gcc tgc ctt cga agt agc cca aag tta aaa aat aag tac ctt 12196  
Leu Ala Ala Cys Leu Arg Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu  
205 210 215

ttt tgg gcc agg tac ttg aca caa ttc caa atg ttc cag ttt atg ctg 12244  
Phe Trp Gly Arg Tyr Leu Thr Gln Phe Gln Met Phe Gln Phe Met Leu

220	225	230	
aac tta gtg cag gct tac tac gac atg aaa acg aat gcg cca tat cca			12292
Asn Leu Val Gln Ala Tyr Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro			
235	240	245	250
caa tgg ctg atc aag att ttg ttc tac tac atg atc tcg ttg ctg ttt			12340
Gln Trp Leu Ile Lys Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe			
	255	260	265
ctt ttc ggc aat ttt tac gta caa aaa tac atc aaa ccc tct gac gga			12388
Leu Phe Gly Asn Phe Tyr Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly			
	270	275	280
aag caa aag gga gct aaa act gag tga tctagaaggc ctctgcttt			12435
Lys Gln Lys Gly Ala Lys Thr Glu			
	285	290	
aatgagatat gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg			12495
ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa			12555
tgaatatatc acccgttact atcgtatttt tatgaataat attctccgtt caatttactg			12615
attgtccgtc gagcaaattt acacattgcc actaaacgtc taaacccttg taatttgttt			12675
ttgttttact atgtgtgtta tgtatttgat ttgcgataaa tttttatatt tggactaaa			12735
tttataacac cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt			12795
gattctaaat tatttttgct ttctaaatac atatactaata caactggaaa tgtaaataatt			12855
tgctaataatt tctactatag gagaattaaa gtgagtgaat atggtaccac aagggttgga			12915
gatttaattg ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga			12975
ggataataat ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaagggtt			13035
agtaattttt caagacaaca atgttaccac acacaagttt tgaggtgcat gcatggatgc			13095
cctgtggaaa gtttaaaaat attttggaat tgatttgcat ggaagccatg tgtaaaacca			13155
tgacatccac ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta			13215
tgcatgtagt ctatataatg aggattttgc aatactttca ttcatacaca ctactaagt			13275
tttacacgat tataatttct tcatagccag cggatcc atg gta ttc gcg ggc ggt			13330
		Met Val Phe Ala Gly Gly	
		295	
gga ctt cag cag ggc tct ctc gaa gaa aac atc gac gtc gag cac att			13378
Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn Ile Asp Val Glu His Ile			
	300	305	310
gcc agt atg tct ctc ttc agc gac ttc ttc agt tat gtg tct tca act			13426
Ala Ser Met Ser Leu Phe Ser Asp Phe Phe Ser Tyr Val Ser Ser Thr			
	315	320	325

## 100

gtt ggt tcg tgg agc gta cac agt ata caa cct ttg aag cgc ctg acg Val Gly Ser Trp Ser Val His Ser Ile Gln Pro Leu Lys Arg Leu Thr 330 335 340 345	13474
agt aag aag cgt gtt tcg gaa agc gct gcc gtg caa tgt ata tca gct Ser Lys Lys Arg Val Ser Glu Ser Ala Ala Val Gln Cys Ile Ser Ala 350 355 360	13522
gaa gtt cag aga aat tcg agt acc cag gga act gcg gag gca ctc gca Glu Val Gln Arg Asn Ser Ser Thr Gln Gly Thr Ala Glu Ala Leu Ala 365 370 375	13570
gaa tca gtc gtg aag ccc acg aga cga agg tca tct cag tgg aag aag Glu Ser Val Val Lys Pro Thr Arg Arg Arg Ser Ser Gln Trp Lys Lys 380 385 390	13618
tcg aca cac ccc cta tca gaa gta gca gta cac aac aag cca agc gat Ser Thr His Pro Leu Ser Glu Val Ala Val His Asn Lys Pro Ser Asp 395 400 405	13666
tgc tgg att gtt gta aaa aac aag gtg tat gat gtt tcc aat ttt gcg Cys Trp Ile Val Val Lys Asn Lys Val Tyr Asp Val Ser Asn Phe Ala 410 415 420 425	13714
gac gag cat ccc gga gga tca gtt att agt act tat ttt gga cga gac Asp Glu His Pro Gly Gly Ser Val Ile Ser Thr Tyr Phe Gly Arg Asp 430 435 440	13762
ggc aca gat gtt ttc tct agt ttt cat gca gct tct aca tgg aaa att Gly Thr Asp Val Phe Ser Ser Phe His Ala Ala Ser Thr Trp Lys Ile 445 450 455	13810
ctt caa gac ttt tac att ggt gac gtg gag agg gtg gag ccg act cca Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu Arg Val Glu Pro Thr Pro 460 465 470	13858
gag ctg ctg aaa gat ttc cga gaa atg aga gct ctt ttc ctg agg gag Glu Leu Leu Lys Asp Phe Arg Glu Met Arg Ala Leu Phe Leu Arg Glu 475 480 485	13906
caa ctt ttc aaa agt tcg aaa ttg tac tat gtt atg aag ctg ctc acg Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr Val Met Lys Leu Leu Thr 490 495 500 505	13954
aat gtt gct att ttt gct gcg agc att gca ata ata tgt tgg agc aag Asn Val Ala Ile Phe Ala Ala Ser Ile Ala Ile Ile Cys Trp Ser Lys 510 515 520	14002
act att tca gcg gtt ttg gct tca gct tgt atg atg gct ctg tgt ttc Thr Ile Ser Ala Val Leu Ala Ser Ala Cys Met Met Ala Leu Cys Phe 525 530 535	14050
caa cag tgc gga tgg cta tcc cat gat ttt ctc cac aat cag gtg ttt Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His Asn Gln Val Phe 540 545 550	14098
gag aca cgc tgg ctt aat gaa gtt gtc ggg tat gtg atc ggc aac gcc Glu Thr Arg Trp Leu Asn Glu Val Val Gly Tyr Val Ile Gly Asn Ala	14146

555	560	565	
ggt ctg ggg ttt agt aca ggg tgg tgg aag gag aag cat aac ctt cat Val Leu Gly Phe Ser Thr Gly Trp Trp Lys Glu Lys His Asn Leu His 570 575 580 585			14194
cat gct gct cca aat gaa tgc gat cag act tac caa cca att gat gaa His Ala Ala Pro Asn Glu Cys Asp Gln Thr Tyr Gln Pro Ile Asp Glu 590 595 600			14242
gat att gat act ctc ccc ctc att gcc tgg agc aag gac ata ctg gcc Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp Ser Lys Asp Ile Leu Ala 605 610 615			14290
aca gtt gag aat aag aca ttc ttg cga atc ctc caa tac cag cat ctg Thr Val Glu Asn Lys Thr Phe Leu Arg Ile Leu Gln Tyr Gln His Leu 620 625 630			14338
ttc ttc atg ggt ctg tta ttt ttc gcc cgt ggt agt tgg ctc ttt tgg Phe Phe Met Gly Leu Leu Phe Phe Ala Arg Gly Ser Trp Leu Phe Trp 635 640 645			14386
agc tgg aga tat acc tct aca gca gtg ctc tca cct gtc gac agg ttg Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu Ser Pro Val Asp Arg Leu 650 655 660 665			14434
ttg gag aag gga act gtt ctg ttt cac tac ttt tgg ttc gtc ggg aca Leu Glu Lys Gly Thr Val Leu Phe His Tyr Phe Trp Phe Val Gly Thr 670 675 680			14482
gcg tgc tat ctt ctc cct ggt tgg aag cca tta gta tgg atg gcg gtg Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro Leu Val Trp Met Ala Val 685 690 695			14530
act gag ctc atg tcc ggc atg ctg ctg gcc ttt gta ttt gta ctt agc Thr Glu Leu Met Ser Gly Met Leu Leu Gly Phe Val Phe Val Leu Ser 700 705 710			14578
cac aat ggg atg gag gtt tat aat tcc tct aaa gaa ttc gtg agt gca His Asn Gly Met Glu Val Tyr Asn Ser Ser Lys Glu Phe Val Ser Ala 715 720 725			14626
cag atc gta tcc aca cgg gat atc aaa gga aac ata ttc aac gac tgg Gln Ile Val Ser Thr Arg Asp Ile Lys Gly Asn Ile Phe Asn Asp Trp 730 735 740 745			14674
ttc act ggt ggc ctt aac agg caa ata gag cat cat ctt ttc cca aca Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro Thr 750 755 760			14722
atg ccc agg cat aat tta aac aaa ata gca cct aga gtg gag gtg ttc Met Pro Arg His Asn Leu Asn Lys Ile Ala Pro Arg Val Glu Val Phe 765 770 775			14770
tgt aag aaa cac ggt ctg gtg tac gaa gac gta tct att gct acc ggc Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Ile Ala Thr Gly 780 785 790			14818

act tgc aag gtt ttg aaa gca ttg aag gaa gtc gcg gag gct gcg gca 14866  
 Thr Cys Lys Val Leu Lys Ala Leu Lys Glu Val Ala Glu Ala Ala Ala  
 795 800 805

gag cag cat gct acc acc agt taa gctagcgta accctgcttt aatgagatat 14920  
 Glu Gln His Ala Thr Thr Ser  
 810 815

gcgagacgcc tatgatcgca tgatatttgc tttcaattct gttgtgcacg ttgtaaaaaa 14980  
 cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc 15040  
 acccgttact atcgtaaaaa tatgaataat attctccggt caatttactg attgtccgctc 15100  
 gagcaaattt acacattgcc actaaacgct taaacccttg taatttggtt ttgttttact 15160  
 atgtgtgta tgtatttgat ttgcgataaa tttttatatt tggactaaa tttataacac 15220  
 cttttatgct aacgtttgcc aacacttagc aatttgcaag ttgattaatt gattctaaat 15280  
 tatttttgtc ttctaaatac atatactaata caactggaaa tgtaaataatt tgctaataatt 15340  
 tctactatag gagaattaaa gtgagtgaat atgggtaccac aagggttgga gatttaattg 15400  
 ttgcaatgct gcatggatgg catatacacc aaacattcaa taattcttga ggataataat 15460  
 ggtaccacac aagatttgag gtgcatgaac gtcacgtgga caaaagggtt agtaattttt 15520  
 caagacaaca atgttaccac acacaagttt tgagggtgcat gcatggatgc cctgtggaaa 15580  
 gtttaaaaat attttggaat tgatttgcac ggaagccatg tgtaaaacca tgacatccac 15640  
 ttggaggatg caataatgaa gaaaactaca aatttacatg caactagtta tgcatgtagt 15700  
 ctatataatg aggattttgc aatactttca ttcatacaca ctactaagt tttacacgat 15760  
 tataatttct tcatagccag cagatctaaa atg gct ccg gat gcg gat aag ctt 15814  
 Met Ala Pro Asp Ala Asp Lys Leu  
 820 825

cga caa cgc cag acg act gcg gta gcg aag cac aat gct gct acc ata 15862  
 Arg Gln Arg Gln Thr Thr Ala Val Ala Lys His Asn Ala Ala Thr Ile  
 830 835 840

tog acg cag gaa cgc ctt tgc agt ctg tct tgc ctc aaa ggc gaa gaa 15910  
 Ser Thr Gln Glu Arg Leu Cys Ser Leu Ser Ser Leu Lys Gly Glu Glu  
 845 850 855

gtc tgc atc gac gga atc atc tat gac ctc caa tca ttc gat cat ccc 15958  
 Val Cys Ile Asp Gly Ile Ile Tyr Asp Leu Gln Ser Phe Asp His Pro  
 860 865 870

ggg ggt gaa acg atc aaa atg ttt ggt ggc aac gat gtc act gta cag 16006  
 Gly Gly Glu Thr Ile Lys Met Phe Gly Gly Asn Asp Val Thr Val Gln  
 875 880 885

tac aag atg att cac ccg tac cat acc gag aag cat ttg gaa aag atg 16054  
 Tyr Lys Met Ile His Pro Tyr His Thr Glu Lys His Leu Glu Lys Met

890	895	900	905	
aag cgt gtc ggc aag gtg acg gat ttc gtc tgc gag tac aag ttc gat				16102
Lys Arg Val Gly Lys Val Thr Asp Phe Val Cys Glu Tyr Lys Phe Asp	910	915	920	
acc gaa ttt gaa cgc gaa atc aaa cga gaa gtc ttc aag att gtg cga				16150
Thr Glu Phe Glu Arg Glu Ile Lys Arg Glu Val Phe Lys Ile Val Arg	925	930	935	
cga ggc aag gat ttc ggt act ttg gga tgg ttc ttc cgt gcg ttt tgc				16198
Arg Gly Lys Asp Phe Gly Thr Leu Gly Trp Phe Phe Arg Ala Phe Cys	940	945	950	
tac att gcc att ttc ttc tac ctg cag tac cat tgg gtc acc acg gga				16246
Tyr Ile Ala Ile Phe Phe Tyr Leu Gln Tyr His Trp Val Thr Thr Gly	955	960	965	
acc tct tgg ctg ctg gcc gtg gcc tac gga atc tcc caa gcg atg att				16294
Thr Ser Trp Leu Leu Ala Val Ala Tyr Gly Ile Ser Gln Ala Met Ile	970	975	980	985
ggc atg aat gtc cag cac gat gcc aac cac ggg gcc acc tcc aag cgt				16342
Gly Met Asn Val Gln His Asp Ala Asn His Gly Ala Thr Ser Lys Arg	990	995	1000	
ccc tgg gtc aac gac atg cta ggc ctc ggt gcg gat ttt att ggt ggt				16390
Pro Trp Val Asn Asp Met Leu Gly Leu Gly Ala Asp Phe Ile Gly Gly	1005	1010	1015	
tcc aag tgg ctc tgg cag gaa caa cac tgg acc cac cac gct tac acc				16438
Ser Lys Trp Leu Trp Gln Glu Gln His Trp Thr His His Ala Tyr Thr	1020	1025	1030	
aat cac gcc gag atg gat ccc gat agc ttt ggt gcc gaa cca atg ctc				16486
Asn His Ala Glu Met Asp Pro Asp Ser Phe Gly Ala Glu Pro Met Leu	1035	1040	1045	
cta ttc aac gac tat ccc ttg gat cat ccc gct cgt acc tgg cta cat				16534
Leu Phe Asn Asp Tyr Pro Leu Asp His Pro Ala Arg Thr Trp Leu His	1050	1055	1060	1065
cgc ttt caa gca ttc ttt tac atg ccc gtc ttg gct gga tac tgg ttg				16582
Arg Phe Gln Ala Phe Phe Tyr Met Pro Val Leu Ala Gly Tyr Trp Leu	1070	1075	1080	
tcc gct gtc ttc aat cca caa att ctt gac ctc cag caa cgc ggc gca				16630
Ser Ala Val Phe Asn Pro Gln Ile Leu Asp Leu Gln Gln Arg Gly Ala	1085	1090	1095	
ctt tcc gtc ggt atc cgt ctc gac aac gct ttc att cac tcg cga cgc				16678
Leu Ser Val Gly Ile Arg Leu Asp Asn Ala Phe Ile His Ser Arg Arg	1100	1105	1110	
aag tat gcg gtt ttc tgg cgg gct gtg tac att gcg gtg aac gtg att				16726
Lys Tyr Ala Val Phe Trp Arg Ala Val Tyr Ile Ala Val Asn Val Ile	1115	1120	1125	

gct ccg ttt tac aca aac tcc ggc ctc gaa tgg tcc tgg cgt gtc ttt 16774  
 Ala Pro Phe Tyr Thr Asn Ser Gly Leu Glu Trp Ser Trp Arg Val Phe  
 1130 1135 1140 1145  
 gga aac atc atg ctc atg ggt gtg gcg gaa tcg ctc gcg ctg gcg gtc 16822  
 Gly Asn Ile Met Leu Met Gly Val Ala Glu Ser Leu Ala Leu Ala Val  
 1150 1155 1160  
 ctg ttt tcg ttg tcg cac aat ttc gaa tcc gcg gat cgc gat ccg acc 16870  
 Leu Phe Ser Leu Ser His Asn Phe Glu Ser Ala Asp Arg Asp Pro Thr  
 1165 1170 1175  
 gcc cca ctg aaa aag acg gga gaa cca gtc gac tgg ttc aag aca cag 16918  
 Ala Pro Leu Lys Lys Thr Gly Glu Pro Val Asp Trp Phe Lys Thr Gln  
 1180 1185 1190  
 gtc gaa act tcc tgc act tac ggt gga ttc ctt tcc ggt tgc ttc acg 16966  
 Val Glu Thr Ser Cys Thr Tyr Gly Gly Phe Leu Ser Gly Cys Phe Thr  
 1195 1200 1205  
 gga ggt ctc aac ttt cag gtt gaa cac cac ttg ttc cca cgc atg agc 17014  
 Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe Pro Arg Met Ser  
 1210 1215 1220 1225  
 agc gct tgg tat ccc tac att gcc ccc aag gtc cgc gaa att tgc gcc 17062  
 Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Lys Val Arg Glu Ile Cys Ala  
 1230 1235 1240  
 aaa cac ggc gtc cac tac gcc tac tac ccg tgg atc cac caa aac ttt 17110  
 Lys His Gly Val His Tyr Ala Tyr Tyr Pro Trp Ile His Gln Asn Phe  
 1245 1250 1255  
 ctc tcc acc gtc cgc tac atg cac gcg gcc ggg acc ggt gcc aac tgg 17158  
 Leu Ser Thr Val Arg Tyr Met His Ala Ala Gly Thr Gly Ala Asn Trp  
 1260 1265 1270  
 cgc cag atg gcc aga gaa aat ccc ttg acc gga cgg gcg taa 17200  
 Arg Gln Met Ala Arg Glu Asn Pro Leu Thr Gly Arg Ala  
 1275 1280 1285  
 agatctgccg gcacgatcc cgggccatgg cctgctttaa tgagatatgc gagacgccta 17260  
 tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg 17320  
 tagctcagat ccttaccgcc gggttcgggt cattctaattg aatatatcac ccgttactat 17380  
 cgtattttta tgaataatat tctccgttca atttactgat tgtccgtcga cgagctcggc 17440  
 gcgcctctag aggatcgatg aattcagatc ggctgagtgg ctcccttcaac gttgcgggttc 17500  
 tgtcagttcc aaacgtaaaa cggcttgtcc cgcgtcatcg gcgggggtca taacgtgact 17560  
 cccttaattc tccgctcatg atcagattgt cgtttccgc cttcagttta aactatcagt 17620  
 gtttgacagg atatattggc gggtaaacct aagagaaaag agcgtttatt agaataatcg 17680  
 gatatttaaa agggcgtgaa aagggttatc cttcgtccat ttgtatgtgc atgccaacca 17740

cagggttccc ca

&lt;210&gt; 29

&lt;211&gt; 290

&lt;212&gt; PRT

&lt;213&gt; Unknown

&lt;400&gt; 29

Met Glu Val Val Glu Arg Phe Tyr Gly Glu Leu Asp Gly Lys Val Ser  
 1 5 10 15

Gln Gly Val Asn Ala Leu Leu Gly Ser Phe Gly Val Glu Leu Thr Asp  
 20 25 30

Thr Pro Thr Thr Lys Gly Leu Pro Leu Val Asp Ser Pro Thr Pro Ile  
 35 40 45

Val Leu Gly Val Ser Val Tyr Leu Thr Ile Val Ile Gly Gly Leu Leu  
 50 55 60

Trp Ile Lys Ala Arg Asp Leu Lys Pro Arg Ala Ser Glu Pro Phe Leu  
 65 70 75 80

Leu Gln Ala Leu Val Leu Val His Asn Leu Phe Cys Phe Ala Leu Ser  
 85 90 95

Leu Tyr Met Cys Val Gly Ile Ala Tyr Gln Ala Ile Thr Trp Arg Tyr  
 100 105 110

Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys His Lys Glu Met Ala Ile  
 115 120 125

Leu Val Tyr Leu Phe Tyr Met Ser Lys Tyr Val Glu Phe Met Asp Thr  
 130 135 140

Val Ile Met Ile Leu Lys Arg Ser Thr Arg Gln Ile Ser Phe Leu His  
 145 150 155 160

Val Tyr His His Ser Ser Ile Ser Leu Ile Trp Trp Ala Ile Ala His  
 165 170 175

His Ala Pro Gly Gly Glu Ala Tyr Trp Ser Ala Ala Leu Asn Ser Gly  
 180 185 190

Val His Val Leu Met Tyr Ala Tyr Tyr Phe Leu Ala Ala Cys Leu Arg  
 195 200 205

Ser Ser Pro Lys Leu Lys Asn Lys Tyr Leu Phe Trp Gly Arg Tyr Leu  
 210 215 220

Thr Gln Phe Gln Met Phe Gln Phe Met Leu Asn Leu Val Gln Ala Tyr  
 225 230 235 240

Tyr Asp Met Lys Thr Asn Ala Pro Tyr Pro Gln Trp Leu Ile Lys Ile  
 245 250 255

Leu Phe Tyr Tyr Met Ile Ser Leu Leu Phe Leu Phe Gly Asn Phe Tyr

260                      265                      270  
 Val Gln Lys Tyr Ile Lys Pro Ser Asp Gly Lys Gln Lys Gly Ala Lys  
       275                      280                      285  
 Thr Glu  
       290  
  
 <210> 30  
 <211> 525  
 <212> PRT  
 <213> Unknown  
  
 <400> 30  
 Met Val Phe Ala Gly Gly Gly Leu Gln Gln Gly Ser Leu Glu Glu Asn  
       1                      5                      10                      15  
 Ile Asp Val Glu His Ile Ala Ser Met Ser Leu Phe Ser Asp Phe Phe  
                   20                      25                      30  
 Ser Tyr Val Ser Ser Thr Val Gly Ser Trp Ser Val His Ser Ile Gln  
                   35                      40                      45  
 Pro Leu Lys Arg Leu Thr Ser Lys Lys Arg Val Ser Glu Ser Ala Ala  
                   50                      55                      60  
 Val Gln Cys Ile Ser Ala Glu Val Gln Arg Asn Ser Ser Thr Gln Gly  
       65                      70                      75                      80  
 Thr Ala Glu Ala Leu Ala Glu Ser Val Val Lys Pro Thr Arg Arg Arg  
                   85                      90                      95  
 Ser Ser Gln Trp Lys Lys Ser Thr His Pro Leu Ser Glu Val Ala Val  
                   100                      105                      110  
 His Asn Lys Pro Ser Asp Cys Trp Ile Val Val Lys Asn Lys Val Tyr  
                   115                      120                      125  
 Asp Val Ser Asn Phe Ala Asp Glu His Pro Gly Gly Ser Val Ile Ser  
                   130                      135                      140  
 Thr Tyr Phe Gly Arg Asp Gly Thr Asp Val Phe Ser Ser Phe His Ala  
       145                      150                      155                      160  
 Ala Ser Thr Trp Lys Ile Leu Gln Asp Phe Tyr Ile Gly Asp Val Glu  
                   165                      170                      175  
 Arg Val Glu Pro Thr Pro Glu Leu Leu Lys Asp Phe Arg Glu Met Arg  
                   180                      185                      190  
 Ala Leu Phe Leu Arg Glu Gln Leu Phe Lys Ser Ser Lys Leu Tyr Tyr  
                   195                      200                      205  
 Val Met Lys Leu Leu Thr Asn Val Ala Ile Phe Ala Ala Ser Ile Ala  
                   210                      215                      220  
 Ile Ile Cys Trp Ser Lys Thr Ile Ser Ala Val Leu Ala Ser Ala Cys

225		230		235		240
Met Met Ala Leu Cys Phe Gln Gln Cys Gly Trp Leu Ser His Asp Phe						
	245			250		255
Leu His Asn Gln Val Phe Glu Thr Arg Trp Leu Asn Glu Val Val Gly						
	260			265		270
Tyr Val Ile Gly Asn Ala Val Leu Gly Phe Ser Thr Gly Trp Trp Lys						
	275			280		285
Glu Lys His Asn Leu His His Ala Ala Pro Asn Glu Cys Asp Gln Thr						
	290			295		300
Tyr Gln Pro Ile Asp Glu Asp Ile Asp Thr Leu Pro Leu Ile Ala Trp						
305		310		315		320
Ser Lys Asp Ile Leu Ala Thr Val Glu Asn Lys Thr Phe Leu Arg Ile						
	325			330		335
Leu Gln Tyr Gln His Leu Phe Phe Met Gly Leu Leu Phe Phe Ala Arg						
	340			345		350
Gly Ser Trp Leu Phe Trp Ser Trp Arg Tyr Thr Ser Thr Ala Val Leu						
	355			360		365
Ser Pro Val Asp Arg Leu Leu Glu Lys Gly Thr Val Leu Phe His Tyr						
	370			375		380
Phe Trp Phe Val Gly Thr Ala Cys Tyr Leu Leu Pro Gly Trp Lys Pro						
385		390		395		400
Leu Val Trp Met Ala Val Thr Glu Leu Met Ser Gly Met Leu Leu Gly						
	405			410		415
Phe Val Phe Val Leu Ser His Asn Gly Met Glu Val Tyr Asn Ser Ser						
	420			425		430
Lys Glu Phe Val Ser Ala Gln Ile Val Ser Thr Arg Asp Ile Lys Gly						
	435			440		445
Asn Ile Phe Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu						
	450			455		460
His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Lys Ile Ala						
	465			470		475
Pro Arg Val Glu Val Phe Cys Lys Lys His Gly Leu Val Tyr Glu Asp						
	485			490		495
Val Ser Ile Ala Thr Gly Thr Cys Lys Val Leu Lys Ala Leu Lys Glu						
	500			505		510
Val Ala Glu Ala Ala Ala Glu Gln His Ala Thr Thr Ser						
	515			520		525

<211> 469  
 <212> PRT  
 <213> Unknown

<400> 31

Met	Ala	Pro	Asp	Ala	Asp	Lys	Leu	Arg	Gln	Arg	Gln	Thr	Thr	Ala	Val
1				5					10					15	
Ala	Lys	His	Asn	Ala	Ala	Thr	Ile	Ser	Thr	Gln	Glu	Arg	Leu	Cys	Ser
			20					25					30		
Leu	Ser	Ser	Leu	Lys	Gly	Glu	Glu	Val	Cys	Ile	Asp	Gly	Ile	Ile	Tyr
	35					40						45			
Asp	Leu	Gln	Ser	Phe	Asp	His	Pro	Gly	Gly	Glu	Thr	Ile	Lys	Met	Phe
	50					55					60				
Gly	Gly	Asn	Asp	Val	Thr	Val	Gln	Tyr	Lys	Met	Ile	His	Pro	Tyr	His
65					70					75				80	
Thr	Glu	Lys	His	Leu	Glu	Lys	Met	Lys	Arg	Val	Gly	Lys	Val	Thr	Asp
				85					90					95	
Phe	Val	Cys	Glu	Tyr	Lys	Phe	Asp	Thr	Glu	Phe	Glu	Arg	Glu	Ile	Lys
			100					105					110		
Arg	Glu	Val	Phe	Lys	Ile	Val	Arg	Arg	Gly	Lys	Asp	Phe	Gly	Thr	Leu
	115						120					125			
Gly	Trp	Phe	Phe	Arg	Ala	Phe	Cys	Tyr	Ile	Ala	Ile	Phe	Phe	Tyr	Leu
	130					135					140				
Gln	Tyr	His	Trp	Val	Thr	Thr	Gly	Thr	Ser	Trp	Leu	Leu	Ala	Val	Ala
145					150					155					160
Tyr	Gly	Ile	Ser	Gln	Ala	Met	Ile	Gly	Met	Asn	Val	Gln	His	Asp	Ala
				165					170					175	
Asn	His	Gly	Ala	Thr	Ser	Lys	Arg	Pro	Trp	Val	Asn	Asp	Met	Leu	Gly
			180					185					190		
Leu	Gly	Ala	Asp	Phe	Ile	Gly	Gly	Ser	Lys	Trp	Leu	Trp	Gln	Glu	Gln
	195						200					205			
His	Trp	Thr	His	His	Ala	Tyr	Thr	Asn	His	Ala	Glu	Met	Asp	Pro	Asp
	210					215					220				
Ser	Phe	Gly	Ala	Glu	Pro	Met	Leu	Leu	Phe	Asn	Asp	Tyr	Pro	Leu	Asp
225					230					235				240	
His	Pro	Ala	Arg	Thr	Trp	Leu	His	Arg	Phe	Gln	Ala	Phe	Phe	Tyr	Met
			245						250					255	
Pro	Val	Leu	Ala	Gly	Tyr	Trp	Leu	Ser	Ala	Val	Phe	Asn	Pro	Gln	Ile
		260					265						270		
Leu	Asp	Leu	Gln	Gln	Arg	Gly	Ala	Leu	Ser	Val	Gly	Ile	Arg	Leu	Asp
	275						280					285			

Asn Ala Phe Ile His Ser Arg Arg Lys Tyr Ala Val Phe Trp Arg Ala  
290 295 300

Val Tyr Ile Ala Val Asn Val Ile Ala Pro Phe Tyr Thr Asn Ser Gly  
305 310 315 320

Leu Glu Trp Ser Trp Arg Val Phe Gly Asn Ile Met Leu Met Gly Val  
325 330 335

Ala Glu Ser Leu Ala Leu Ala Val Leu Phe Ser Leu Ser His Asn Phe  
340 345 350

Glu Ser Ala Asp Arg Asp Pro Thr Ala Pro Leu Lys Lys Thr Gly Glu  
355 360 365

Pro Val Asp Trp Phe Lys Thr Gln Val Glu Thr Ser Cys Thr Tyr Gly  
370 375 380

Gly Phe Leu Ser Gly Cys Phe Thr Gly Gly Leu Asn Phe Gln Val Glu  
385 390 395 400

His His Leu Phe Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala  
405 410 415

Pro Lys Val Arg Glu Ile Cys Ala Lys His Gly Val His Tyr Ala Tyr  
420 425 430

Tyr Pro Trp Ile His Gln Asn Phe Leu Ser Thr Val Arg Tyr Met His  
435 440 445

Ala Ala Gly Thr Gly Ala Asn Trp Arg Gln Met Ala Arg Glu Asn Pro  
450 455 460

Leu Thr Gly Arg Ala  
465

<210> 32

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Polylinker

<400> 32

gaattcggcg cgccgagctc ctcgag

26

<210> 33

<211> 265

<212> DNA

<213> Artificial Sequence

<220>

<223> Polylinker-Terminator-Polylinker

<400> 33  
ccaccgcggt gggcgccgc ctgcagtcta gaaggcctcc tgctttaatg agatatgcga 60  
gacgcctatg atcgcatgat atttgctttc aattctgttg tgcacgttgt aaaaaacctg 120  
agcatgtgta gtcagatcc ttaccgcccg tttcggttca ttctaataa tatatcacc 180  
gttactatcg ttttttatg aataatattc tccgttcaat ttactgattg tccgtcgacg 240  
aattcgagct cggcgccca agctt 265

<210> 34  
<211> 257  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Polylinker-Terminator-Polylinker

<400> 34  
ggatccgata tggggccgc tagcgtaac cctgctttaa tgagatatgc gagacgccta 60  
tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg 120  
tagctcagat ccttaccgcc ggtttcgggt cattctaataa aatataatcac ccgttactat 180  
cgtattttta tgaataatat tctccgttca atttactgat tgtccgtcga cgaattcgag 240  
ctcggcgccg caagctt 257

<210> 35  
<211> 257  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Polylinker-Terminator-Polylinker

<400> 35  
agatctgccg gcatcgatcc cgggccatgg cctgctttaa tgagatatgc gagacgccta 60  
tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc tgagcatgtg 120  
tagctcagat ccttaccgcc ggtttcgggt cattctaataa aatataatcac ccgttactat 180  
cgtattttta tgaataatat tctccgttca atttactgat tgtccgtcga cgaattcgag 240  
ctcggcgccg caagctt 257